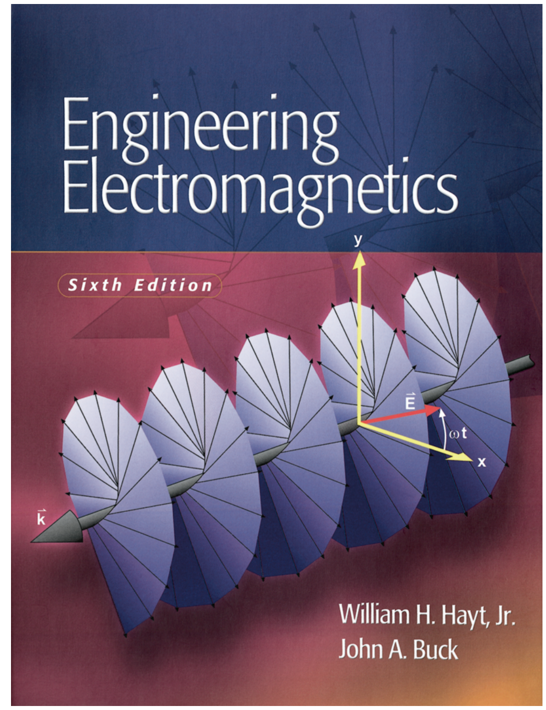
Interactive e-Text

? Help | Feedback



Engineering Electromagnetics

Sixth Edition

William H. Hayt, Jr. · John A. Buck

Textbook Table of Contents

The Textbook Table of Contents is your starting point for accessing pages within the chapter. Once you're at this location, you can easily move back and forth within specific chapters or just as easily jump from one chapter to another.

Textbook Website

The Textbook Website is the McGraw-Hill Higher Education website developed to accompany this textbook. Here you'll find numerous text-specific learning tools and resources that expand upon the information you normally find in a printed textbook.

McGraw-Hill Website

The McGraw-Hill Website is your starting point for discovery of all the educational content and services offered by McGraw-Hill Higher Education.

Copyright @ 2001 The McGraw Companies. All rights reserved. Any use is subject to the Terms of Use and Privacy Policy. McGraw-Hill Higher Education is one of the many fine businesses of The McGraw-Hill Companies.

If you have a question or a suggestion about a specific book or product, please fill out our User Feedback Form accessible from the main menu or contact our customer service line at 1-800-262-4729.

The McGraw-Hill Companies

Engineering Electromagnetics

McGraw-Hill Series in Electrical and Computer Engineering

SENIOR CONSULTING EDITOR

Stephen W. Director, University of Michigan, Ann Arbor

Circuits and Systems
Communications and Signal Processing
Computer Engineering
Control Theory and Robotics
Electromagnetics
Electronics and VLSI Circuits
Introductory
Power
Antennas, Microwaves, and Radar

Previous Consulting Editors

Ronald N. Bracewell, Colin Cherry, James F. Gibbons, Willis W. Harman, Hubert Heffner, Edward W. Herold, John G. Linvill, Simon Ramo, Ronald A. Rohrer, Anthony E. Siegman, Charles Susskind, Frederick E. Terman, John G. Truxal, Ernst Weber, and John R. Whinnery

Engineering Electromagnetics

SIXTH EDITION

William H. Hayt, Jr.

Late Emeritus Professor
Purdue University

John A. Buck

Georgia Institute of Technology

Boston Burr Ridge, IL Dubuque, IA Madison, WI
New York San Francisco St. Louis
Bangkok Bogotá Caracas Lisbon London Madrid Mexico City
Milan New Delhi Seoul Singapore Sydney Taipei Toronto

BRIEF CONTENTS

	Preface	xi
Chapter 1	**Vector Analysis**	1
Chapter 2	**Coulomb's Law and Electric Field Intensity**	27
Chapter 3	**Electric Flux Density, Gauss' Law, and Divergence**	53
Chapter 4	**Energy and Potential**	83
Chapter 5	**Conductors, Dielectrics, and Capacitance**	119
Chapter 6	**Experimental Mapping Methods**	169
Chapter 7	**Poisson's and Laplace's Equations**	195
Chapter 8	**The Steady Magnetic Field**	224
Chapter 9	**Magnetic Forces, Materials, and Inductance**	274
Chapter 10	**Time-Varying Fields and Maxwell's Equations**	322
Chapter 11	**The Uniform Plane Wave**	348
Chapter 12	**Plane Waves at Boundaries and in Dispersive Media**	387
Chapter 13	**Transmission Lines**	435
Chapter 14	**Waveguide and Antenna Fundamentals**	484
	Appendix A Vector Analysis	529
	Appendix B Units	534
	Appendix C Material Constants	540
	Appendix D Origins of the Complex Permittivity	544
	Appendix E Answers to Selected Problems*	
	Index	551

*To find Appendix E, please visit the expanded book website: www.mhhe.com/engcs/electrical/haytbuck

PREFACE

Over the years, I have developed a familiarity with this book in its various editions, having learned from it, referred to it, and taught from it. The second edition was used in my first electromagnetics course as a junior during the early '70's. Its simple and easy-to-read style convinced me that this material could be learned, and it helped to confirm my latent belief at the time that my specialty would lie in this direction. Later, it was not surprising to see my own students coming to me with heavily-marked copies, asking for help on the drill problems, and taking a more active interest in the subject than I usually observed. So, when approached to be the new co-author, and asked what I would do to change the book, my initial feeling was—nothing. Further reflection brought to mind earlier wishes for more material on waves and transmission lines. As a result, Chapters 1 to 10 are original, while 11 to 14 have been revised, and contain new material.

A conversation with Bill Hayt at the project's beginning promised the start of what I thought would be a good working relationship. The rapport was immediate. His declining health prevented his active participation, but we seemed to be in general agreement on the approach to a revision. Although I barely knew him, his death, occurring a short time later, deeply affected me in the sense that someone that I greatly respected was gone, along with the promise of a good friendship. My approach to the revision has been as if he were still here. In the front of my mind was the wish to write and incorporate the new material in a manner that he would have approved, and which would have been consistent with the original objectives and theme of the text. Much more could have been done, but at the risk of losing the book's identity and possibly its appeal.

Before their deaths, Bill Hayt and Jack Kemmerly completed an entirely new set of drill problems and end-of-chapter problems for the existing material at that time, up to and including the transmission lines chapter. These have been incorporated, along with my own problems that pertain to the new topics. The other revisions are summarized as follows: The original chapter on plane waves has now become two. The first (Chapter 11) is concerned with the development of the uniform plane wave and the treatment wave propagation in various media. These include lossy materials, where propagation and loss are now modeled in a general way using the complex permittivity. Conductive media are presented as special cases, as are materials that exhibit electronic or molecular resonances. A new appendix provides background on resonant media. A new section on wave polarization is also included. Chapter 12 deals with wave reflection at single and multiple interfaces, and at oblique incidence angles. An additional section on dispersive media has been added, which introduces the concepts of group velocity and group dispersion. The effect of pulse broadening arising from group dispersion is treated at an elementary level. Chapter 13 is essentially the old transmission lines chapter, but with a new section on transients. Chapter 14 is intended as an introduction to waveguides and antennas, in which the underlying

physical concepts are emphasized. The waveguide sections are all new, but the antennas treatment is that of the previous editions.

The approach taken in the new material, as was true in the original work, is to emphasize physical understanding and problem-solving skills. I have also moved the work more in the direction of communications-oriented material, as this seemed a logical way in which the book could evolve, given the material that was already there. The perspective has been broadened by an expanded emphasis toward optics concepts and applications, which are presented along with the more traditional lower-frequency discussions. This again seemed to be a logical step, as the importance of optics and optical communications has increased significantly since the earlier editions were published.

The theme of the text has not changed since the first edition of 1958. An inductive approach is used that is consistent with the historical development. In it, the experimental laws are presented as individual concepts that are later unified in Maxwell's equations. Apart from the first chapter on vector analysis, the mathematical tools are introduced in the text on an as-needed basis. Throughout every edition, as well as this one, the primary goal has been to enable students to learn independently. Numerous examples, drill problems (usually having multiple parts), and end-of-chapter problems are provided to facilitate this. Answers to the drill problems are given below each problem. Answers to selected end-of-chapter problems can be found on the internet at www.mhhe.com/engcs/electrical/haytbuck. A solutions manual is also available.

The book contains more than enough material for a one-semester course. As is evident, statics concepts are emphasized and occur first in the presentation. In a course that places more emphasis on dynamics, the later chapters can be reached earlier by omitting some or all of the material in Chapters 6 and 7, as well as the later sections of Chapter 8. The transmission line treatment (Chapter 13) relies heavily on the plane wave development in Chapters 11 and 12. A more streamlined presentation of plane waves, leading to an earlier arrival at transmission lines, can be accomplished by omitting sections 11.5, 12.5, and 12.6. Chapter 14 is intended as an "advanced topics" chapter, in which the development of waveguide and antenna concepts occurs through the application of the methods learned in earlier chapters, thus helping to solidify that knowledge. It may also serve as a bridge between the basic course and more advanced courses that follow it.

I am deeply indebted to several people who provided much-needed feedback and assistance on the work. Glenn S. Smith, Georgia Tech, reviewed parts of the manuscript and had many suggestions on the content and the philosophy of the revision. Several outside reviewers pointed out errors and had excellent suggestions for improving the presentation, most of which, within time limitations, were taken. These include Madeleine Andrawis, South Dakota State University, M. Yousif El-Ibiary, University of Oklahoma, Joel T. Johnson, Ohio State University, David Kelley, Pennsylvania State University, Sharad R. Laxpati, University of Illinois at Chicago, Masoud Mostafavi, San Jose State University, Vladimir A. Rakov, University of Florida, Hussain Al-Rizzo, Sultan

Qaboos University, Juri Silmberg, Ryerson Polytechnic University and Robert M. Weikle II, University of Virginia. My editors at McGraw-Hill, Catherine Fields, Michelle Flomenhoft, and Betsy Jones, provided excellent expertise and support—particularly Michelle, who was almost in daily contact, and provided immediate and knowledgeable answers to all questions and concerns. My seemingly odd conception of the cover illustration was brought into reality through the graphics talents of Ms Diana Fouts at Georgia Tech. Finally, much is owed to my wife and daughters for putting up with a part-time husband and father for many a weekend.

John A. Buck
Atlanta, 2000

CONTENTS

Preface xi

Chapter 1 Vector Analysis 1

1.1. Scalars and Vectors 2
1.2. Vector Algebra 3
1.3. The Cartesian Coordinate System 4
1.4. Vector Components and Unit Vectors 6
1.5. The Vector Field 9
1.6. The Dot Product 10
1.7. The Cross Product 13
1.8. Other Coordinate Systems: Circular Cylindrical Coordinates 15
1.9. The Spherical Coordinate System 20

Chapter 2 Coulomb's Law and Electric Field Intensity 27

2.1. The Experimental Law of Coulomb 28
2.2. Electric Field Intensity 31
2.3. Field Due to a Continuous Volume Charge Distribution 36
2.4. Field of a Line Charge 38
2.5. Field of a Sheet Charge 44
2.6. Streamlines and Sketches of Fields 46

Chapter 3 Electric Flux Density, Gauss' Law, and Divergence 53

3.1. Electric Flux Density 54
3.2. Gauss' Law 57
3.3. Applications of Gauss' Law: Some Symmetrical Charge Distributions 62
3.4. Application of Gauss' Law: Differential Volume Element 67
3.5. Divergence 70
3.6. Maxwell's First Equation (Electrostatics) 73
3.7. The Vector Operator ∇ and the Divergence Theorem 74

Chapter 4 Energy and Potential 83

4.1. Energy and Potential in a Moving Point Charge in an Electric Field 84
4.2. The Line Integral 85
4.3. Definition of Potential Difference and Potential 91
4.4. The Potential Field of a Point Charge 93

4.5. The Potential Field of a System of Charges: Conservative Property 95
4.6. Potential Gradient 99
4.7. The Dipole 106
4.8. Energy Density in the Electric Field 110

Chapter 5 Conductors, Dielectrics, and Capacitance 119
5.1. Current and Current Density 120
5.2. Continuity of Current 122
5.3. Metallic Conductors 124
5.4. Conductor Properties and Boundary Conditions 129
5.5. The Method of Images 134
5.6. Semiconductors 136
5.7. The Nature of Dielectric Materials 138
5.8. Boundary Conditions for Perfect Dielectric Materials 144
5.9. Capacitance 150
5.10. Several Capacitance Examples 154
5.11. Capacitance of a Two-Wire Line 157

Chapter 6 Experimental Mapping Methods 169
6.1. Curvilinear Squares 170
6.2. The Iteration Method 176
6.3. Current Analogies 183
6.4. Physical Models 186

Chapter 7 Poisson's and Laplace's Equations 195
7.1 Poisson's and Laplace's Equations 196
7.2. Uniqueness Theorem 198
7.3. Examples of the Solution of Laplace's Equation 200
7.4. Example of the Solution of Poisson's Equation 207
7.5. Product Solution of Laplace's Equation 211

Chapter 8 The Steady Magnetic Field 224
8.1. Biot-Savart Law 225
8.2. Ampere's Circuital Law 232
8.3. Curl 239
8.4. Stokes' Theorem 246
8.5. Magnetic Flux and Magnetic Flux Density 251
8.6. The Scalar and Vector Magnetic Potentials 254
8.7. Derivation of the Steady-Magnetic-Field Laws 261

Chapter 9	**Magnetic Forces, Materials and Inductance**	274
	9.1. Force on a Moving Charge	275
	9.2. Force on a Differential Current Element	276
	9.3. Force Between Differential Current Elements	280
	9.4. Force and Torque on a Closed Circuit	283
	9.5. The Nature of Magnetic Materials	288
	9.6. Magnetization and Permeability	292
	9.7. Magnetic Boundary Conditions	297
	9.8. The Magnetic Circuit	299
	9.9. Potential Energy and Forces on Magnetic Materials	306
	9.10. Inductance and Mutual Inductance	308
Chapter 10	**Time-Varying Fields and Maxwell's Equations**	322
	10.1. Faraday's Law	323
	10.2. Displacement Current	329
	10.3. Maxwell's Equations in Point Form	334
	10.4. Maxwell's Equations in Integral Form	336
	10.5. The Retarded Potentials	338
Chapter 11	**The Uniform Plane Wave**	348
	11.1. Wave Propagation in Free Space	348
	11.2. Wave Propagation in Dielectrics	356
	11.3. The Poynting Vector and Power Considerations	365
	11.4. Propagation in Good Conductors: Skin Effect	369
	11.5. Wave Polarization	376
Chapter 12	**Plane Waves at Boundaries and in Dispersive Media**	387
	12.1. Reflection of Uniform Plane Waves at Normal Incidence	388
	12.2. Standing Wave Ratio	395
	12.3. Wave Reflection from Multiple Interfaces	400
	12.4. Plane Wave Propagation in General Directions	408
	12.5. Plane Wave Reflection at Oblique Incidence Angles	411
	12.6. Wave Propagation in Dispersive Media	421
Chapter 13	**Transmission Lines**	435
	13.1. The Transmission-Line Equations	436
	13.2. Transmission-Line Parameters	442
	13.3. Some Transmission-Line Examples	448
	13.4. Graphical Methods	452
	13.5. Several Practical Problems	460
	13.6. Transients on Transmission Lines	463

Chapter 14 **Waveguide and Antenna Fundamentals** 484

14.1. Basic Waveguide Operation 485
14.2. Plane Wave Analysis of the Parallel-Plate Waveguide 488
14.3. Parallel-Plate Guide Analysis Using the Wave Equation 497
14.4. Rectangular Waveguides 501
14.5. Dielectric Waveguides 506
14.6. Basic Antenna Principles 514

Appendix A **Vector Analysis** 529

Appendix B **Units** 534

Appendix C **Material Constants** 540

Appendix D **Origins of the Complex Permittivity** 544

Appendix E **Answers to Selected Problems***

Index 551

*To find Appendix E, please visit the expanded website: www.mhhe.com/engcs/electrical/haytbuck

CHAPTER 1

VECTOR ANALYSIS

Vector analysis is a mathematical subject which is much better taught by mathematicians than by engineers. Most junior and senior engineering students, however, have not had the time (or perhaps the inclination) to take a course in vector analysis, although it is likely that many elementary vector concepts and operations were introduced in the calculus sequence. These fundamental concepts and operations are covered in this chapter, and the time devoted to them now should depend on past exposure.

The viewpoint here is also that of the engineer or physicist and not that of the mathematician in that proofs are indicated rather than rigorously expounded and the physical interpretation is stressed. It is easier for engineers to take a more rigorous and complete course in the mathematics department after they have been presented with a few physical pictures and applications.

It is possible to study electricity and magnetism without the use of vector analysis, and some engineering students may have done so in a previous electrical engineering or basic physics course. Carrying this elementary work a bit further, however, soon leads to line-filling equations often composed of terms which all look about the same. A quick glance at one of these long equations discloses little of the physical nature of the equation and may even lead to slighting an old friend.

Vector analysis is a mathematical shorthand. It has some new symbols, some new rules, and a pitfall here and there like most new fields, and it demands concentration, attention, and practice. The drill problems, first met at the end of Sec. 1.4, should be considered an integral part of the text and should all be

worked. They should not prove to be difficult if the material in the accompanying section of the text has been thoroughly understood. It take a little longer to "read" the chapter this way, but the investment in time will produce a surprising interest.

1.1 SCALARS AND VECTORS

The term *scalar* refers to a quantity whose value may be represented by a single (positive or negative) real number. The x, y, and z we used in basic algebra are scalars, and the quantities they represent are scalars. If we speak of a body falling a distance L in a time t, or the temperature T at any point in a bowl of soup whose coordinates are x, y, and z, then L, t, T, x, y, and z are all scalars. Other scalar quantities are mass, density, pressure (but not force), volume, and volume resistivity. Voltage is also a scalar quantity, although the complex representation of a sinusoidal voltage, an artificial procedure, produces a *complex scalar*, or *phasor*, which requires two real numbers for its representation, such as amplitude and phase angle, or real part and imaginary part.

A *vector* quantity has both a magnitude[1] and a direction in space. We shall be concerned with two- and three-dimensional spaces only, but vectors may be defined in n-dimensional space in more advanced applications. Force, velocity, acceleration, and a straight line from the positive to the negative terminal of a storage battery are examples of vectors. Each quantity is characterized by both a magnitude and a direction.

We shall be mostly concerned with scalar and vector *fields*. A field (scalar or vector) may be defined mathematically as some function of that vector which connects an arbitrary origin to a general point in space. We usually find it possible to associate some physical effect with a field, such as the force on a compass needle in the earth's magnetic field, or the movement of smoke particles in the field defined by the vector velocity of air in some region of space. Note that the field concept invariably is related to a region. Some quantity is defined at every point in a region. Both *scalar fields* and *vector fields* exist. The temperature throughout the bowl of soup and the density at any point in the earth are examples of scalar fields. The gravitational and magnetic fields of the earth, the voltage gradient in a cable, and the temperature gradient in a soldering-iron tip are examples of vector fields. The value of a field varies in general with both position and time.

In this book, as in most others using vector notation, vectors will be indicated by boldface type, for example, **A**. Scalars are printed in italic type, for example, A. When writing longhand or using a typewriter, it is customary to draw a line or an arrow over a vector quantity to show its vector character. (CAUTION: This is the first pitfall. Sloppy notation, such as the omission of the line or arrow symbol for a vector, is the major cause of errors in vector analysis.)

---

[1] We adopt the convention that "magnitude" infers "absolute value"; the magnitude of any quantity is therefore always positive.

1.2 VECTOR ALGEBRA

With the definitions of vectors and vector fields now accomplished, we may proceed to define the rules of vector arithmetic, vector algebra, and (later) of vector calculus. Some of the rules will be similar to those of scalar algebra, some will differ slightly, and some will be entirely new and strange. This is to be expected, for a vector represents more information than does a scalar, and the multiplication of two vectors, for example, will be more involved than the multiplication of two scalars.

The rules are those of a branch of mathematics which is firmly established. Everyone "plays by the same rules," and we, of course, are merely going to look at and interpret these rules. However, it is enlightening to consider ourselves pioneers in the field. We are making our own rules, and we can make any rules we wish. The only requirement is that the rules be self-consistent. Of course, it would be nice if the rules agreed with those of scalar algebra where possible, and it would be even nicer if the rules enabled us to solve a few practical problems.

One should not fall into the trap of "algebra worship" and believe that the rules of college algebra were delivered unto man at the Creation. These rules are merely self-consistent and extremely useful. There are other less familiar algebras, however, with very different rules. In Boolean algebra the product AB can be only unity or zero. Vector algebra has its own set of rules, and we must be constantly on guard against the mental forces exerted by the more familiar rules or scalar algebra.

Vectorial addition follows the parallelogram law, and this is easily, if inaccurately, accomplished graphically. Fig. 1.1 shows the sum of two vectors, $\mathbf{A}$ and $\mathbf{B}$. It is easily seen that $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$, or that vector addition obeys the commutative law. Vector addition also obeys the associative law,

$$\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$$

Note that when a vector is drawn as an arrow of finite length, its location is defined to be at the tail end of the arrow.

Coplanar vectors, or vectors lying in a common plane, such as those shown in Fig. 1.1, which both lie in the plane of the paper, may also be added by expressing each vector in terms of "horizontal" and "vertical" components and adding the corresponding components.

Vectors in three dimensions may likewise be added by expressing the vectors in terms of three components and adding the corresponding components. Examples of this process of addition will be given after vector components are discussed in Sec. 1.4.

The rule for the subtraction of vectors follows easily from that for addition, for we may always express $\mathbf{A}-\mathbf{B}$ as $\mathbf{A}+(-\mathbf{B})$; the sign, or direction, of the second vector is reversed, and this vector is then added to the first by the rule for vector addition.

Vectors may be multiplied by scalars. The magnitude of the vector changes, but its direction does not when the scalar is positive, although it reverses direc-



FIGURE 1.1
Two vectors may be added graphically either by drawing both vectors from a common origin and completing the parallelogram or by beginning the second vector from the head of the first and completing the triangle; either method is easily extended to three or more vectors.

tion when multiplied by a negative scalar. Multiplication of a vector by a scalar also obeys the associative and distributive laws of algebra, leading to

$$(r+s)(\mathbf{A}+\mathbf{B}) = r(\mathbf{A}+\mathbf{B}) + s(\mathbf{A}+\mathbf{B}) = r\mathbf{A} + r\mathbf{B} + s\mathbf{A} + s\mathbf{B}$$

Division of a vector by a scalar is merely multiplication by the reciprocal of that scalar.

The multiplication of a vector by a vector is discussed in Secs. 1.6 and 1.7.

Two vectors are said to be equal if their difference is zero, or $\mathbf{A} = \mathbf{B}$ if $\mathbf{A} - \mathbf{B} = 0$.

In our use of vector fields we shall always add and subtract vectors which are defined at the same point. For example, the *total* magnetic field about a small horseshoe magnet will be shown to be the sum of the fields produced by the earth and the permanent magnet; the total field at any point is the sum of the individual fields at that point.

If we are not considering a vector *field*, however, we may add or subtract vectors which are not defined at the same point. For example, the sum of the gravitational force acting on a 150-lb_f (pound-force) man at the North Pole and that acting on a 175-lb_f man at the South Pole may be obtained by shifting each force vector to the South Pole before addition. The resultant is a force of 25 lb_f directed toward the center of the earth at the South Pole; if we wanted to be difficult, we could just as well describe the force as 25 lb_f directed *away* from the center of the earth (or ‘‘upward’’) at the North Pole.[2]

1.3 THE CARTESIAN COORDINATE SYSTEM

In order to describe a vector accurately, some specific lengths, directions, angles, projections, or components must be given. There are three simple methods of doing this, and about eight or ten other methods which are useful in very special cases. We are going to use only the three simple methods, and the simplest of these is the *cartesian*, or *rectangular*, *coordinate system*.

[2] A few students have argued that the force might be described at the equator as being in a ‘‘northerly’’ direction. They are right, but enough is enough.

In the cartesian coordinate system we set up three coordinate axes mutually at right angles to each other, and call them the x, y, and z axes. It is customary to choose a *right-handed* coordinate system, in which a rotation (through the smaller angle) of the x axis into the y axis would cause a right-handed screw to progress in the direction of the z axis. If the right hand is used, then the thumb, forefinger, and middle finger may then be identified, respectively, as the x, y, and z axes. Fig. 1.2*a* shows a right-handed cartesian coordinate system.

A point is located by giving its x, y, and z coordinates. These are, respectively, the distances from the origin to the intersection of a perpendicular dropped from the point to the x, y, and z axes. An alternative method of interpreting coordinate values, and a method corresponding to that which *must* be used in all other coordinate systems, is to consider the point as being at the

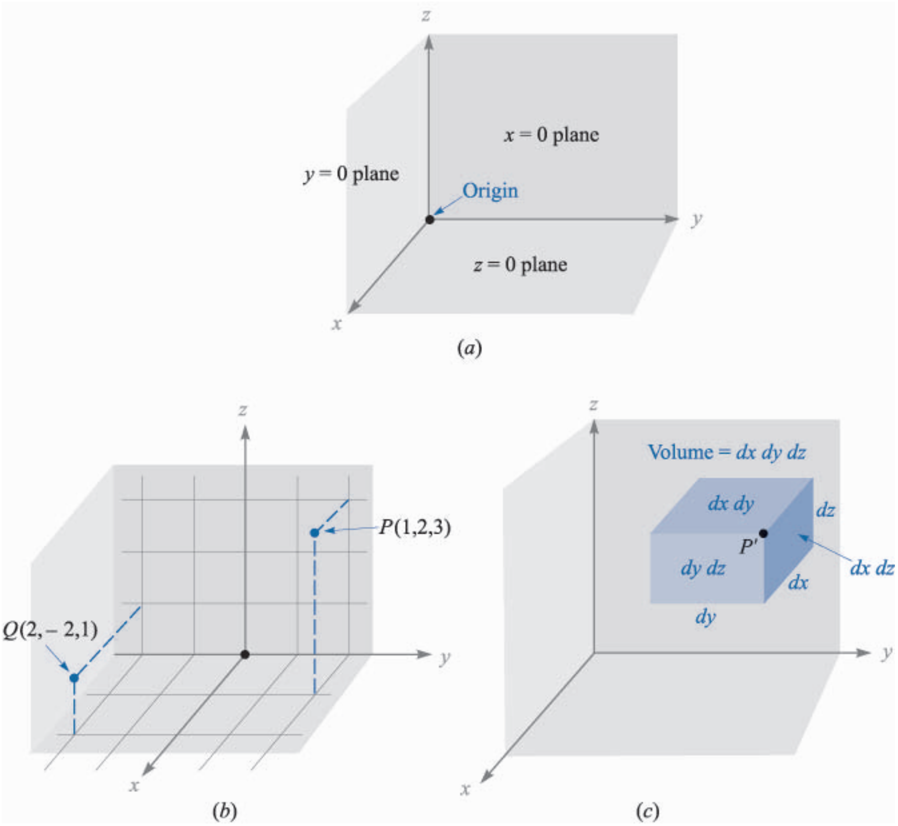


FIGURE 1.2
(*a*) A right-handed cartesian coordinate system. If the curved fingers of the right hand indicate the direction through which the x axis is turned into coincidence with the y axis, the thumb shows the direction of the z axis. (*b*) The location of points $P(1, 2, 3)$ and $Q(2, -2, 1)$. (*c*) The differential volume element in cartesian coordinates; dx, dy, and dz are, in general, independent differentials.

common intersection of three surfaces, the planes $x =$ constant, $y =$ constant, and $z =$ constant, the constants being the coordinate values of the point.

Fig. 1.2*b* shows the points P and Q whose coordinates are $(1, 2, 3)$ and $(2, -2, 1)$, respectively. Point P is therefore located at the common point of intersection of the planes $x = 1$, $y = 2$, and $z = 3$, while point Q is located at the intersection of the planes $x = 2$, $y = -2$, $z = 1$.

As we encounter other coordinate systems in Secs. 1.8 and 1.9, we should expect points to be located at the common intersection of three surfaces, not necessarily planes, but still mutually perpendicular at the point of intersection.

If we visualize three planes intersecting at the general point P, whose coordinates are x, y, and z, we may increase each coordinate value by a differential amount and obtain three slightly displaced planes intersecting at point P', whose coordinates are $x + dx$, $y + dy$, and $z + dz$. The six planes define a rectangular parallelepiped whose volume is $dv = dxdydz$; the surfaces have differential areas dS of $dxdy$, $dydz$, and $dzdx$. Finally, the distance dL from P to P' is the diagonal of the parallelepiped and has a length of $\sqrt{(dx)^2 + (dy)^2 + (dz)^2}$. The volume element is shown in Fig. 1.2*c*; point P' is indicated, but point P is located at the only invisible corner.

All this is familiar from trigonometry or solid geometry and as yet involves only scalar quantities. We shall begin to describe vectors in terms of a coordinate system in the next section.

1.4 VECTOR COMPONENTS AND UNIT VECTORS

To describe a vector in the cartesian coordinate system, let us first consider a vector $\mathbf{r}$ extending outward from the origin. A logical way to identify this vector is by giving the three *component vectors*, lying along the three coordinate axes, whose vector sum must be the given vector. If the component vectors of the vector $\mathbf{r}$ are $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$, then $\mathbf{r} = \mathbf{x} + \mathbf{y} + \mathbf{z}$. The component vectors are shown in Fig. 1.3*a*. Instead of one vector, we now have three, but this is a step forward, because the three vectors are of a very simple nature; each is always directed along one of the coordinate axes.

In other words, the component vectors have magnitudes which depend on the given vector (such as $\mathbf{r}$ above), but they each have a known and constant direction. This suggests the use of *unit vectors* having unit magnitude, by definition, and directed along the coordinate axes in the direction of the increasing coordinate values. We shall reserve the symbol $\mathbf{a}$ for a unit vector and identify the direction of the unit vector by an appropriate subscript. Thus $\mathbf{a}_x$, $\mathbf{a}_y$, and $\mathbf{a}_z$ are the unit vectors in the cartesian coordinate system.[3] They are directed along the x, y, and z axes, respectively, as shown in Fig. 1.3*b*.

[3]The symbols **i**, **j**, and **k** are also commonly used for the unit vectors in cartesian coordinates.

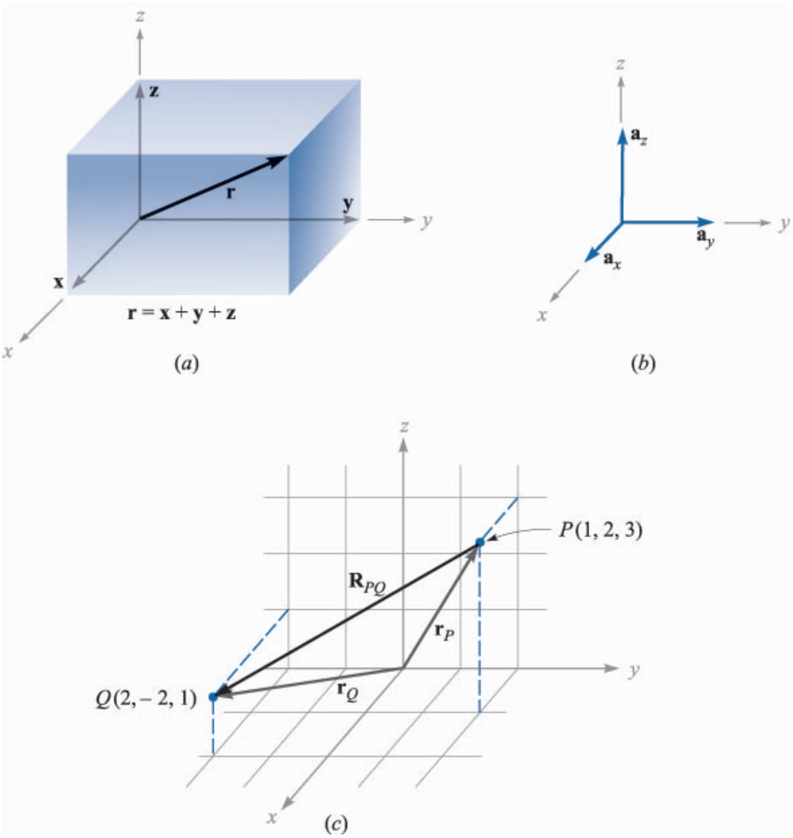


FIGURE 1.3
(*a*) The component vectors **x**, **y**, and **z** of vector **r**. (*b*) The unit vectors of the cartesian coordinate system have unit magnitude and are directed toward increasing values of their respective variables. (*c*) The vector $\mathbf{R}_{PQ}$ is equal to the vector difference $\mathbf{r}_Q - \mathbf{r}_P$.

If the component vector **y** happens to be two units in magnitude and directed toward increasing values of y, we should then write $\mathbf{y} = 2\mathbf{a}_y$. A vector $\mathbf{r}_P$ pointing from the origin to point $P(1, 2, 3)$ is written $\mathbf{r}_P = \mathbf{a}_x + 2\mathbf{a}_y + 3\mathbf{a}_z$. The vector from P to Q may be obtained by applying the rule of vector addition. This rule shows that the vector from the origin to P plus the vector from P to Q is equal to the vector from the origin to Q. The desired vector from $P(1, 2, 3)$ to $Q(2, -2, 1)$ is therefore

$$\mathbf{R}_{PQ} = \mathbf{r}_Q - \mathbf{r}_P = (2-1)\mathbf{a}_x + (-2-2)\mathbf{a}_y + (1-3)\mathbf{a}_z$$
$$= \mathbf{a}_x - 4\mathbf{a}_y - 2\mathbf{a}_z$$

The vectors $\mathbf{r}_P$, $\mathbf{r}_Q$, and $\mathbf{R}_{PQ}$ are shown in Fig. 1.3*c*.

This last vector does not extend outward from the origin, as did the vector **r** we initially considered. However, we have already learned that vectors having the same magnitude and pointing in the same direction are equal, so we see that to help our visualization processes we are at liberty to slide any vector over to the origin before determining its component vectors. Parallelism must, of course, be maintained during the sliding process.

If we are discussing a force vector **F**, or indeed any vector other than a displacement-type vector such as **r**, the problem arises of providing suitable letters for the three component vectors. It would not do to call them **x**, **y**, and **z**, for these are displacements, or directed distances, and are measured in meters (abbreviated m) or some other unit of length. The problem is most often avoided by using *component scalars*, simply called *components*, F_x, F_y, and F_z. The components are the signed magnitudes of the component vectors. We may then write $\mathbf{F} = F_x\mathbf{a}_x + F_y\mathbf{a}_y + F_z\mathbf{a}_z$. The component vectors are $F_x\mathbf{a}_x$, $F_y\mathbf{a}_y$, and $F_z\mathbf{a}_z$.

Any vector **B** then may be described by $\mathbf{B} = B_x\mathbf{a}_x + B_y\mathbf{a}_y + B_z\mathbf{a}_z$. The magnitude of **B** written $|\mathbf{B}|$ or simply B, is given by

$$|\mathbf{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2} \tag{1}$$

Each of the three coordinate systems we discuss will have its three fundamental and mutually perpendicular unit vectors which are used to resolve any vector into its component vectors. However, unit vectors are not limited to this application. It is often helpful to be able to write a unit vector having a specified direction. This is simply done, for a unit vector in a given direction is merely a vector in that direction divided by its magnitude. A unit vector in the **r** direction is $\mathbf{r}/\sqrt{x^2 + y^2 + z^2}$, and a unit vector in the direction of the vector **B** is

$$\mathbf{a}_B = \frac{\mathbf{B}}{\sqrt{B_x^2 + B_y^2 + B_z^2}} = \frac{\mathbf{B}}{|\mathbf{B}|} \tag{2}$$

Example 1.1

Specify the unit vector extending from the origin toward the point $G(2, -2, -1)$.

Solution. We first construct the vector extending from the origin to point G,

$$\mathbf{G} = 2\mathbf{a}_x - 2\mathbf{a}_y - \mathbf{a}_z$$

We continue by finding the magnitude of **G**,

$$|\mathbf{G}| = \sqrt{(2)^2 + (-2)^2 + (-1)^2} = 3$$

and finally expressing the desired unit vector as the quotient,

$$\mathbf{a}_G = \frac{\mathbf{G}}{|\mathbf{G}|} = \tfrac{2}{3}\mathbf{a}_x - \tfrac{2}{3}\mathbf{a}_y - \tfrac{1}{3}\mathbf{a}_z = 0.667\mathbf{a}_x - 0.667\mathbf{a}_y - 0.333\mathbf{a}_z$$

A special identifying symbol is desirable for a unit vector so that its character is immediately apparent. Symbols which have been used are $\mathbf{u}_B$, $\mathbf{a}_B$, $\mathbf{1_B}$, or even $\mathbf{b}$. We shall consistently use the lowercase **a** with an appropriate subscript.

[NOTE: Throughout the text, drill problems appear following sections in which a new principle is introduced in order to allow students to test their understanding of the basic fact itself. The problems are useful in gaining familiarization with new terms and ideas and should all be worked. More general problems appear at the ends of the chapters. The answers to the drill problems are given in the same order as the parts of the problem.]

✔ **D1.1.** Given points $M(-1, 2, 1)$, $N(3, -3, 0)$, and $P(-2, -3, -4)$, find: (*a*) $\mathbf{R}_{MN}$; (*b*) $\mathbf{R}_{MN} + \mathbf{R}_{MP}$; (*c*) $|\mathbf{r}_M|$; (*d*) $\mathbf{a}_{MP}$; (*e*) $|2\mathbf{r}_P - 3\mathbf{r}_N|$.

Ans. $4\mathbf{a}_x - 5\mathbf{a}_y - \mathbf{a}_z$; $3\mathbf{a}_x - 10\mathbf{a}_y - 6\mathbf{a}_z$; 2.45; $-0.1400\mathbf{a}_x - 0.700\mathbf{a}_y - 0.700\mathbf{a}_z$; 15.56

1.5 THE VECTOR FIELD

We have already defined a vector field as a vector function of a position vector. In general, the magnitude and direction of the function will change as we move throughout the region, and the value of the vector function must be determined using the coordinate values of the point in question. Since we have considered only the cartesian coordinate system, we should expect the vector to be a function of the variables x, y, and z.

If we again represent the position vector as $\mathbf{r}$, then a vector field $\mathbf{G}$ can be expressed in functional notation as $\mathbf{G}(\mathbf{r})$; a scalar field T is written as $T(\mathbf{r})$.

If we inspect the velocity of the water in the ocean in some region near the surface where tides and currents are important, we might decide to represent it by a velocity vector which is in any direction, even up or down. If the z axis is taken as upward, the x axis in a northerly direction, the y axis to the west, and the origin at the surface, we have a right-handed coordinate system and may write the velocity vector as $\mathbf{v} = v_x\mathbf{a}_x + v_y\mathbf{a}_y + v_z\mathbf{a}_z$, or $\mathbf{v}(\mathbf{r}) = v_x(\mathbf{r})\mathbf{a}_x + v_y(\mathbf{r})\mathbf{a}_y + v_z(\mathbf{r})\mathbf{a}_z$; each of the components v_x, v_y, and v_z may be a function of the three variables x, y, and z. If the problem is simplified by assuming that we are in some portion of the Gulf Stream where the water is moving only to the north, then v_y, and v_z are zero. Further simplifying assumptions might be made if the velocity falls off with depth and changes very slowly as we move north, south, east, or west. A suitable expression could be $\mathbf{v} = 2e^{z/100}\mathbf{a}_x$. We have a velocity of 2 m/s (meters per second) at the surface and a velocity of 0.368×2, or 0.736 m/s, at a depth of 100 m ($z = -100$), and the velocity continues to decrease with depth; in this example the vector velocity has a constant direction.

While the example given above is fairly simple and only a rough approximation to a physical situation, a more exact expression would be correspond-

ingly more complex and difficult to interpret. We shall come across many fields in our study of electricity and magnetism which are simpler than the velocity example, an example in which only the component and one variable were involved (the x component and the variable z). We shall also study more complicated fields, and methods of interpreting these expressions physically will be discussed then.

✔ **D1.2.** A vector field $\mathbf{S}$ is expressed in cartesian coordinates as $\mathbf{S} = \{125/[(x-1)^2 + (y-2)^2 + (z+1)^2]\}\{(x-1)\mathbf{a}_x + (y-2)\mathbf{a}_y + (z+1)\mathbf{a}_z\}$. (*a*) Evaluate $\mathbf{S}$ at $P(2, 4, 3)$. (*b*) Determine a unit vector that gives the direction of $\mathbf{S}$ at P. (*c*) Specify the surface $f(x, y, z)$ on which $|\mathbf{S}| = 1$.

Ans. $5.95\mathbf{a}_x + 11.90\mathbf{a}_y + 23.8\mathbf{a}_z$; $0.218\mathbf{a}_x + 0.436\mathbf{a}_y + 0.873\mathbf{a}_z$; $\sqrt{(x-1)^2 + (y-2)^2 + (z+1)^2} = 125$

1.6 THE DOT PRODUCT

We now consider the first of two types of vector multiplication. The second type will be discussed in the following section.

Given two vectors $\mathbf{A}$ and $\mathbf{B}$, the *dot product*, or *scalar product*, is defined as the product of the magnitude of $\mathbf{A}$, the magnitude of $\mathbf{B}$, and the cosine of the smaller angle between them,

$$\boxed{\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}|\,|\mathbf{B}| \cos\theta_{AB}} \tag{3}$$

The dot appears between the two vectors and should be made heavy for emphasis. The dot, or scalar, product is a scalar, as one of the names implies, and it obeys the commutative law,

$$\boxed{\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}} \tag{4}$$

for the sign of the angle does not affect the cosine term. The expression $\mathbf{A} \cdot \mathbf{B}$ is read "$\mathbf{A}$ dot $\mathbf{B}$."

Perhaps the most common application of the dot product is in mechanics, where a constant force $\mathbf{F}$ applied over a straight displacement $\mathbf{L}$ does an amount of work $FL\cos\theta$, which is more easily written $\mathbf{F} \cdot \mathbf{L}$. We might anticipate one of the results of Chap. 4 by pointing out that if the force varies along the path, integration is necessary to find the total work, and the result becomes

$$\text{Work} = \int \mathbf{F} \cdot d\mathbf{L}$$

Another example might be taken from magnetic fields, a subject about which we shall have a lot more to say later. The total flux Φ crossing a surface

of area S is given by BS if the magnetic flux density B is perpendicular to the surface and uniform over it. We define a *vector surface* $\mathbf{S}$ as having the usual area for its magnitude and having a direction *normal* to the surface (avoiding for the moment the problem of which of the two possible normals to take). The flux crossing the surface is then $\mathbf{B}\cdot\mathbf{S}$. This expression is valid for any direction of the uniform magnetic flux density. However, if the flux density is not constant over the surface, the total flux is $\Phi = \int \mathbf{B}\cdot d\mathbf{S}$. Integrals of this general form appear in Chap. 3 when we study electric flux density.

Finding the angle between two vectors in three-dimensional space is often a job we would prefer to avoid, and for that reason the definition of the dot product is usually not used in its basic form. A more helpful result is obtained by considering two vectors whose cartesian components are given, such as $\mathbf{A} = A_x\mathbf{a}_x + A_y\mathbf{a}_y + A_z\mathbf{a}_z$ and $\mathbf{B} = B_x\mathbf{a}_x + B_y\mathbf{a}_y + B_z\mathbf{a}_z$. The dot product also obeys the distributive law, and, therefore, $\mathbf{A}\cdot\mathbf{B}$ yields the sum of nine scalar terms, each involving the dot product of two unit vectors. Since the angle between two different unit vectors of the cartesian coordinate system is 90°, we then have

$$\mathbf{a}_x\cdot\mathbf{a}_y = \mathbf{a}_y\cdot\mathbf{a}_x = \mathbf{a}_x\cdot\mathbf{a}_z = \mathbf{a}_z\cdot\mathbf{a}_x = \mathbf{a}_y\cdot\mathbf{a}_z = \mathbf{a}_z\cdot\mathbf{a}_y = 0$$

The remaining three terms involve the dot product of a unit vector with itself, which is unity, giving finally

$$\boxed{\mathbf{A}\cdot\mathbf{B} = A_xB_x + A_yB_y + A_zB_z} \tag{5}$$

which is an expression involving no angles.

A vector dotted with itself yields the magnitude squared, or

$$\boxed{\mathbf{A}\cdot\mathbf{A} = A^2 = |\mathbf{A}|^2} \tag{6}$$

and any unit vector dotted with itself is unity,

$$\mathbf{a}_A\cdot\mathbf{a}_A = 1$$

One of the most important applications of the dot product is that of finding the component of a vector in a given direction. Referring to Fig. 1.4*a*, we can obtain the component (scalar) of $\mathbf{B}$ in the direction specified by the unit vector $\mathbf{a}$ as

$$\mathbf{B}\cdot\mathbf{a} = |\mathbf{B}|\,|\mathbf{a}|\cos\theta_{Ba} = |\mathbf{B}|\cos\theta_{Ba}$$

The sign of the component is positive if $0 \le \theta_{Ba} \le 90°$ and negative whenever $90° \le \theta_{Ba} \le 180°$.

In order to obtain the component *vector* of $\mathbf{B}$ in the direction of $\mathbf{a}$, we simply multiply the component (scalar) by $\mathbf{a}$, as illustrated by Fig. 1.4*b*. For example, the component of $\mathbf{B}$ in the direction of $\mathbf{a}_x$ is $\mathbf{B}\cdot\mathbf{a}_x = B_x$, and the

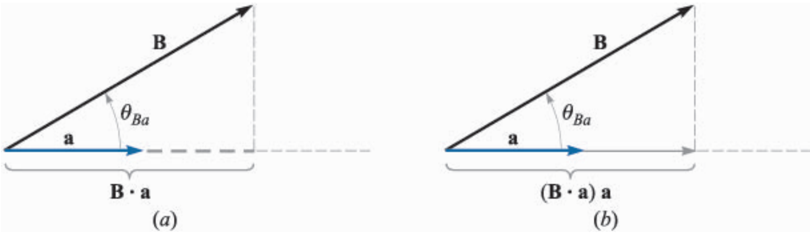


FIGURE 1.4
(*a*) The scalar component of **B** in the direction of the unit vector **a** is $\mathbf{B}\cdot\mathbf{a}$. (*b*) The vector component of **B** in the direction of the unit vector **a** is $(\mathbf{B}\cdot\mathbf{a})\mathbf{a}$.

component vector is $B_x\mathbf{a}_x$, or $(\mathbf{B}\cdot\mathbf{a}_x)\mathbf{a}_x$. Hence, the problem of finding the component of a vector in any desired direction becomes the problem of finding a unit vector in that direction, and that we can do.

The geometrical term *projection* is also used with the dot product. Thus, $\mathbf{B}\cdot\mathbf{a}$ is the projection of **B** in the **a** direction.

Example 1.2

In order to illustrate these definitions and operations, let us consider the vector field $\mathbf{G} = y\mathbf{a}_x - 2.5x\mathbf{a}_y + 3\mathbf{a}_z$ and the point $Q(4, 5, 2)$. We wish to find: **G** at Q; the scalar component of **G** at Q in the direction of $\mathbf{a}_N = \frac{1}{3}(2\mathbf{a}_x + \mathbf{a}_y - 2\mathbf{a}_z)$; the vector component of **G** at Q in the direction of $\mathbf{a}_N$; and finally, the angle θ_{Ga} between $\mathbf{G}(\mathbf{r}_Q)$ and $\mathbf{a}_N$.

Solution. Substituting the coordinates of point Q into the expression for **G**, we have

$$\mathbf{G}(\mathbf{r}_Q) = 5\mathbf{a}_x - 10\mathbf{a}_y + 3\mathbf{a}_z$$

Next we find the scalar component. Using the dot product, we have

$$\mathbf{G}\cdot\mathbf{a}_N = (5\mathbf{a}_x - 10\mathbf{a}_y + 3\mathbf{a}_z)\cdot\tfrac{1}{3}(2\mathbf{a}_x + \mathbf{a}_y - 2\mathbf{a}_z) = \tfrac{1}{3}(10 - 10 - 6) = -2$$

The vector component is obtained by multiplying the scalar component by the unit vector in the direction of $\mathbf{a}_N$,

$$(\mathbf{G}\cdot\mathbf{a}_N)\mathbf{a}_N = -(2)\tfrac{1}{3}(2\mathbf{a}_x + \mathbf{a}_y - 2\mathbf{a}_z) = -1.333\mathbf{a}_x - 0.667\mathbf{a}_y + 1.333\mathbf{a}_z$$

The angle between $\mathbf{G}(\mathbf{r}_Q)$ and $\mathbf{a}_N$ is found from

$$\mathbf{G}\cdot\mathbf{a}_N = |\mathbf{G}|\cos\theta_{Ga}$$
$$-2 = \sqrt{25 + 100 + 9}\cos\theta_{Ga}$$

and

$$\theta_{Ga} = \cos^{-1}\frac{-2}{\sqrt{134}} = 99.9^\circ$$

✔ **D1.3.** The three vertices of a triangle are located at $A(6, -1, 2)$, $B(-2, 3, -4)$, and $C(-3, 1, 5)$. Find: (*a*) $\mathbf{R}_{AB}$; (*b*) $\mathbf{R}_{AC}$; (*c*) the angle θ_{BAC} at vertex A; (*d*) the (vector) projection of $\mathbf{R}_{AB}$ on $\mathbf{R}_{AC}$.

Ans. $-8\mathbf{a}_x + 4\mathbf{a}_y - 6\mathbf{a}_z$; $-9\mathbf{a}_x - 2\mathbf{a}_y + 3\mathbf{a}_z$; 53.6°; $-5.94\mathbf{a}_x + 1.319\mathbf{a}_y + 1.979\mathbf{a}_z$

1.7 THE CROSS PRODUCT

Given two vectors **A** and **B**, we shall now define the *cross product*, or *vector product*, of **A** and **B**, written with a cross between the two vectors as $\mathbf{A} \times \mathbf{B}$ and read "**A** cross **B**." The cross product $\mathbf{A} \times \mathbf{B}$ is a vector; the magnitude of $\mathbf{A} \times \mathbf{B}$ is equal to the product of the magnitudes of **A**, **B**, and the sine of the smaller angle between **A** and **B**; the direction of $\mathbf{A} \times \mathbf{B}$ is perpendicular to the plane containing **A** and **B** and is along that one of the two possible perpendiculars which is in the direction of advance of a right-handed screw as **A** is turned into **B**. This direction is illustrated in Fig. 1.5. Remember that either vector may be moved about at will, maintaining its direction constant, until the two vectors have a "common origin." This determines the plane containing both. However, in most of our applications we shall be concerned with vectors defined at the same point.

As an equation we can write

$$\boxed{\mathbf{A} \times \mathbf{B} = \mathbf{a}_N |\mathbf{A}|\,|\mathbf{B}| \sin \theta_{AB}} \tag{7}$$

where an additional statement, such as that given above, is still required to explain the direction of the unit vector $\mathbf{a}_N$. The subscript stands for "normal."

Reversing the order of the vectors **A** and **B** results in a unit vector in the opposite direction, and we see that the cross product is not commutative, for $\mathbf{B} \times \mathbf{A} = -(\mathbf{A} \times \mathbf{B})$. If the definition of the cross product is applied to the unit

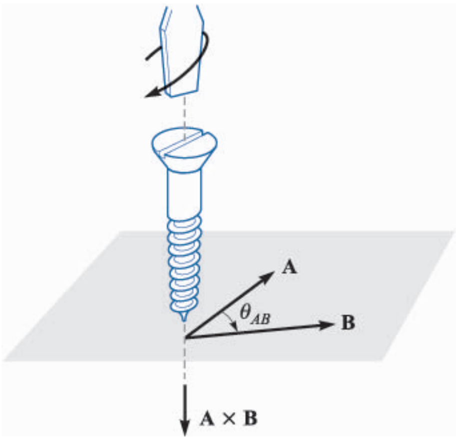


FIGURE 1.5
The direction of $\mathbf{A} \times \mathbf{B}$ is in the direction of advance of a right-handed screw as **A** is turned into **B**.

vectors $\mathbf{a}_x$ and $\mathbf{a}_y$, we find $\mathbf{a}_x \times \mathbf{a}_y = \mathbf{a}_z$, for each vector has unit magnitude, the two vectors are perpendicular, and the rotation of $\mathbf{a}_x$ into $\mathbf{a}_y$ indicates the positive z direction by the definition of a right-handed coordinate system. In a similar way $\mathbf{a}_y \times \mathbf{a}_z = \mathbf{a}_x$, and $\mathbf{a}_z \times \mathbf{a}_x = \mathbf{a}_y$. Note the alphabetic symmetry. As long as the three vectors $\mathbf{a}_x$, $\mathbf{a}_y$, and $\mathbf{a}_z$ are written in order (and assuming that $\mathbf{a}_x$ follows $\mathbf{a}_z$, like three elephants in a circle holding tails, so that we could also write $\mathbf{a}_y$, $\mathbf{a}_z$, $\mathbf{a}_x$ or $\mathbf{a}_z$, $\mathbf{a}_x$, $\mathbf{a}_y$), then the cross and equal sign may be placed in either of the two vacant spaces. As a matter of fact, it is now simpler to define a right-handed cartesian coordinate system by saying that $\mathbf{a}_x \times \mathbf{a}_y = \mathbf{a}_z$.

A simple example of the use of the cross product may be taken from geometry or trigonometry. To find the area of a parallelogram, the product of the lengths of two adjacent sides is multiplied by the sine of the angle between them. Using vector notation for the two sides, we then may express the (scalar) area as the *magnitude* of $\mathbf{A} \times \mathbf{B}$, or $|\mathbf{A} \times \mathbf{B}|$.

The cross product may be used to replace the right-hand rule familiar to all electrical engineers. Consider the force on a straight conductor of length $\mathbf{L}$, where the direction assigned to $\mathbf{L}$ corresponds to the direction of the steady current I, and a uniform magnetic field of flux density $\mathbf{B}$ is present. Using vector notation, we may write the result neatly as $\mathbf{F} = I\mathbf{L} \times \mathbf{B}$. This relationship will be obtained later in Chap. 9.

The evaluation of a cross product by means of its definition turns out to be more work than the evaluation of the dot product from its definition, for not only must we find the angle between the vectors, but we must find an expression for the unit vector $\mathbf{a}_N$. This work may be avoided by using cartesian components for the two vectors $\mathbf{A}$ and $\mathbf{B}$ and expanding the cross product as a sum of nine simpler cross products, each involving two unit vectors,

$$\begin{aligned}\mathbf{A} \times \mathbf{B} = &A_x B_x \mathbf{a}_x \times \mathbf{a}_x + A_x B_y \mathbf{a}_x \times \mathbf{a}_y + A_x B_z \mathbf{a}_x \times \mathbf{a}_z \\ &+ A_y B_x \mathbf{a}_y \times \mathbf{a}_x + A_y B_y \mathbf{a}_y \times \mathbf{a}_y + A_y B_z \mathbf{a}_y \times \mathbf{a}_z \\ &+ A_z B_x \mathbf{a}_z \times \mathbf{a}_x + A_z B_y \mathbf{a}_z \times \mathbf{a}_y + A_z B_z \mathbf{a}_z \times \mathbf{a}_z\end{aligned}$$

We have already found that $\mathbf{a}_x \times \mathbf{a}_y = \mathbf{a}_z$, $\mathbf{a}_y \times \mathbf{a}_z = \mathbf{a}_x$, and $\mathbf{a}_z \times \mathbf{a}_x = \mathbf{a}_y$. The three remaining terms are zero, for the cross product of any vector with itself is zero, since the included angle is zero. These results may be combined to give

$$\mathbf{A} \times \mathbf{B} = (A_y B_z - A_z B_y)\mathbf{a}_x + (A_z B_x - A_x B_z)\mathbf{a}_y + (A_x B_y - A_y B_x)\mathbf{a}_z \tag{8}$$

or written as a determinant in a more easily remembered form,

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{a}_x & \mathbf{a}_y & \mathbf{a}_z \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \tag{9}$$

Thus, if $\mathbf{A} = 2\mathbf{a}_x - 3\mathbf{a}_y + \mathbf{a}_z$ and $\mathbf{B} = -4\mathbf{a}_x - 2\mathbf{a}_y + 5\mathbf{a}_z$, we have

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{a}_x & \mathbf{a}_y & \mathbf{a}_z \\ 2 & -3 & 1 \\ -4 & -2 & 5 \end{vmatrix}$$

$$= [(-3)(5) - (1(-2)]\mathbf{a}_x - [(2)(5) - (1)(-4)]\mathbf{a}_y + [(2)(-2) - (-3)(-4)]\mathbf{a}_z$$

$$= -13\mathbf{a}_x - 14\mathbf{a}_y - 16\mathbf{a}_z$$

✔ **D1.4.** The three vertices of a triangle are located at $A(6, -1, 2)$, $B(-2, 3, -4)$ and $C(-3, 1, 5)$. Find: (*a*) $\mathbf{R}_{AB} \times \mathbf{R}_{AC}$; (*b*) the area of the triangle; (*c*) a unit vector perpendicular to the plane in which the triangle is located.

Ans. $24\mathbf{a}_x + 78\mathbf{a}_y + 20\mathbf{a}_z$; 42.0; $0.286\mathbf{a}_x + 0.928\mathbf{a}_y + 0.238\mathbf{a}_z$

1.8 OTHER COORDINATE SYSTEMS: CIRCULAR CYLINDRICAL COORDINATES

The cartesian coordinate system is generally the one in which students prefer to work every problem. This often means a lot more work for the student, because many problems possess a type of symmetry which pleads for a more logical treatment. It is easier to do now, once and for all, the work required to become familiar with cylindrical and spherical coordinates, instead of applying an equal or greater effort to every problem involving cylindrical or spherical symmetry later. With this future saving of labor in mind, we shall take a careful and unhurried look at cylindrical and spherical coordinates.

The circular cylindrical coordinate system is the three-dimensional version of the polar coordinates of analytic geometry. In the two-dimensional polar coordinates, a point was located in a plane by giving its distance ρ from the origin, and the angle ϕ between the line from the point to the origin and an arbitrary radial line, taken as $\phi = 0$.[4] A three-dimensional coordinate system, circular cylindrical coordinates, is obtained by also specifying the distance z of the point from an arbitrary $z = 0$ reference plane which is perpendicular to the line $\rho = 0$. For simplicity, we usually refer to circular cylindrical coordinates simply as cylindrical coordinates. This will not cause any confusion in reading this book, but it is only fair to point out that there are such systems as elliptic cylindrical coordinates, hyperbolic cylindrical coordinates, parabolic cylindrical coordinates, and others.

We no longer set up three axes as in cartesian coordinates, but must instead consider any point as the intersection of three mutually perpendicular surfaces. These surfaces are a circular cylinder ($\rho =$ constant), a plane ($\phi =$ constant), and

---

[4] The two variables of polar coordinates are commonly called r and θ. With three coordinates, however, it is more common to use ρ for the radius variable of cylindrical coordinates and r for the (different) radius variable of spherical coordinates. Also, the angle variable of cylindrical coordinates is customarily called ϕ because everyone uses θ for a different angle in spherical coordinates. The angle ϕ is common to both cylindrical and spherical coordinates. See?

another plane ($z =$ constant). This corresponds to the location of a point in a cartesian coordinate system by the intersection of three planes ($x =$ constant, $y =$ constant, and $z =$ constant). The three surfaces of circular cylindrical coordinates are shown in Fig. 1.6*a*. Note that three such surfaces may be passed through any point, unless it lies on the z axis, in which case one plane suffices.

Three unit vectors must also be defined, but we may no longer direct them along the "coordinate axes," for such axes exist only in cartesian coordinates. Instead, we take a broader view of the unit vectors in cartesian coordinates and realize that they are directed toward increasing coordinate values and are perpendicular to the surface on which that coordinate value is constant (i.e., the unit vector $\mathbf{a}_x$ is normal to the plane $x =$ constant and points toward larger values of x). In a corresponding way we may now define three unit vectors in cylindrical coordinates, $\mathbf{a}_\rho$, $\mathbf{a}_\phi$, and $\mathbf{a}_z$.

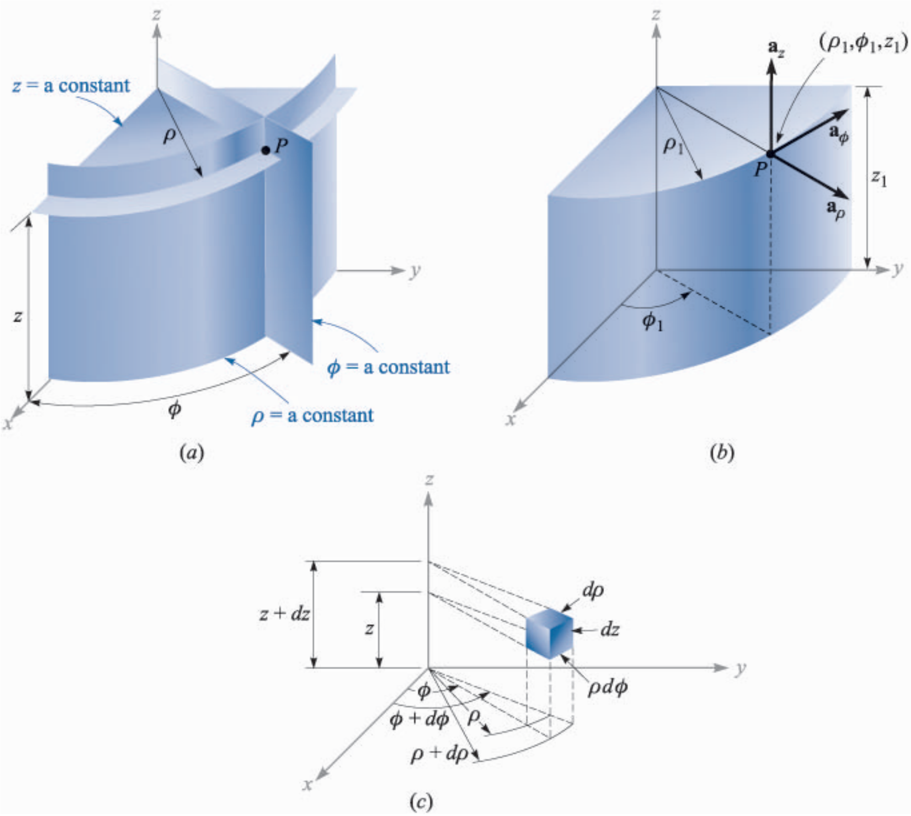


FIGURE 1.6
(a*a*) The three mutually perpendicular surfaces of the circular cylindrical coordinate system. (*b*) The three unit vectors of the circular cylindrical coordinate system. (*c*) The differential volume unit in the circular cylindrical coordinate system; $d\rho$, $\rho d\phi$, and dz are all elements of length.

The unit vector $\mathbf{a}_\rho$ at a point $P(\rho_1, \phi_1, z_1)$ is directed radially outward, normal to the cylindrical surface $\rho = \rho_1$. It lies in the planes $\phi = \phi_1$ and $z = z_1$. The unit vector $\mathbf{a}_\phi$ is normal to the plane $\phi = \phi_1$, points in the direction of increasing ϕ, lies in the plane $z = z_1$, and is tangent to the cylindrical surface $\rho = \rho_1$. The unit vector $\mathbf{a}_z$ is the same as the unit vector $\mathbf{a}_z$ of the cartesian coordinate system. Fig. 1.6*b* shows the three vectors in cylindrical coordinates.

In cartesian coordinates, the unit vectors are not functions of the coordinates. Two of the unit vectors in cylindrical coordinates, $\mathbf{a}_\rho$ and $\mathbf{a}_\phi$, however, *do* vary with the coordinate ϕ, since their directions change. In integration or differentiation with respect to ϕ, then, $\mathbf{a}_\rho$ and $\mathbf{a}_\phi$ must not be treated as constants.

The unit vectors are again mutually perpendicular, for each is normal to one of the three mutually perpendicular surfaces, and we may define a right-handed cylindrical coordinate system as one in which $\mathbf{a}_\rho \times \mathbf{a}_\phi = \mathbf{a}_z$, or (for those who have flexible fingers) as one in which the thumb, forefinger, and middle finger point in the direction of increasing ρ, ϕ, and z, respectively.

A differential volume element in cylindrical coordinates may be obtained by increasing ρ, ϕ, and z by the differential increments $d\rho$, $d\phi$, and dz. The two cylinders of radius ρ and $\rho + d\rho$, the two radial planes at angles ϕ and $\phi + d\phi$, and the two "horizontal" planes at "elevations" z and $z + dz$ now enclose a small volume, as shown in Fig. 1.6*c*, having the shape of a truncated wedge. As the volume element becomes very small, its shape approaches that of a rectangular parallelepiped having sides of length $d\rho$, $\rho d\phi$ and dz. Note that $d\rho$ and dz are dimensionally lengths, but $d\phi$ is not; $\rho d\phi$ is the length. The surfaces have areas of $\rho\, d\rho\, d\phi$, $d\rho\, dz$, and $\rho\, d\phi\, dz$, and the volume becomes $\rho\, d\rho\, d\phi\, dz$.

The variables of the rectangular and cylindrical coordinate systems are easily related to each other. With reference to Fig. 1.7, we see that

$$\begin{aligned} x &= \rho \cos\phi \\ y &= \rho \sin\phi \\ z &= z \end{aligned} \tag{10}$$

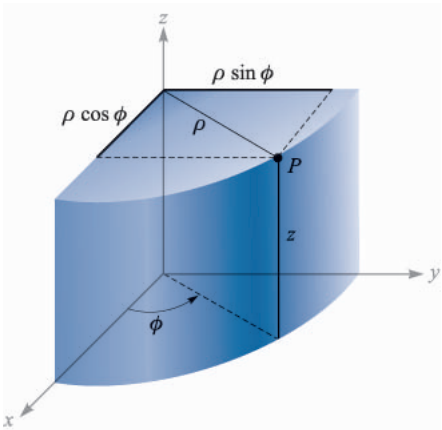


FIGURE 1.7
The relationship between the cartesian variables x, y, z and the cylindrical coordinate variables ρ, ϕ, z. There is no change in the variable z between the two systems.

From the other viewpoint, we may express the cylindrical variables in terms of x, y, and z:

$$\rho = \sqrt{x^2 + y^2} \quad (\rho \geq 0)$$
$$\phi = \tan^{-1}\frac{y}{x} \tag{11}$$
$$z = z$$

We shall consider the variable ρ to be positive or zero, thus using only the positive sign for the radical in (11). The proper value of the angle ϕ is determined by inspecting the signs of x and y. Thus, if $x = -3$ and $y = 4$, we find that the point lies in the second quadrant so that $\rho = 5$ and $\phi = 126.9^\circ$. For $x = 3$ and $y = -4$, we have $\phi = -53.1^\circ$ or 306.9°, whichever is more convenient.

Using (10) or (11), scalar functions given in one coordinate system are easily transformed into the other system.

A vector function in one coordinate system, however, requires two steps in order to transform it to another coordinate system, because a different set of component vectors is generally required. That is, we may be given a cartesian vector

$$\mathbf{A} = A_x\mathbf{a}_x + A_y\mathbf{a}_y + A_z\mathbf{a}_z$$

where each component is given as a function of x, y, and z, and we need a vector in cylindrical coordinates

$$\mathbf{A} = A_\rho\mathbf{a}_\rho + A_\phi\mathbf{a}_\phi + A_z\mathbf{a}_z$$

where each component is given as a function of ρ, ϕ, and z.

To find any desired component of a vector, we recall from the discussion of the dot product that a component in a desired direction may be obtained by taking the dot product of the vector and a unit vector in the desired direction. Hence,

$$A_\rho = \mathbf{A} \cdot \mathbf{a}_\rho \qquad \text{and} \qquad A_\phi = \mathbf{A} \cdot \mathbf{a}_\phi$$

Expanding these dot products, we have

$$A_\rho = (A_x\mathbf{a}_x + A_y\mathbf{a}_y + A_z\mathbf{a}_z) \cdot \mathbf{a}_\rho = A_x\mathbf{a}_x \cdot \mathbf{a}_\rho + A_y\mathbf{a}_y \cdot \mathbf{a}_\rho \tag{12}$$
$$A_\phi = (A_x\mathbf{a}_x + A_y\mathbf{a}_y + A_z\mathbf{a}_z) \cdot \mathbf{a}_\phi = A_x\mathbf{a}_x \cdot \mathbf{a}_\phi + A_y\mathbf{a}_y \cdot \mathbf{a}_\phi \tag{13}$$

and

$$A_z = (A_x\mathbf{a}_x + A_y\mathbf{a}_y + A_z\mathbf{a}_z) \cdot \mathbf{a}_z = A_z\mathbf{a}_z \cdot \mathbf{a}_z = A_z \tag{14}$$

since $\mathbf{a}_z \cdot \mathbf{a}_\rho$ and $\mathbf{a}_z \cdot \mathbf{a}_\phi$ are zero.

In order to complete the transformation of the components, it is necessary to know the dot products $\mathbf{a}_x \cdot \mathbf{a}_\rho$, $\mathbf{a}_y \cdot \mathbf{a}_\rho$, $\mathbf{a}_x \cdot \mathbf{a}_\phi$, and $\mathbf{a}_y \cdot \mathbf{a}_\phi$. Applying the definition of the dot product, we see that since we are concerned with unit vectors, the result is merely the cosine of the angle between the two unit vectors in question. Referring to Fig. 1.7 and thinking mightily, we identify the angle between $\mathbf{a}_x$ and

TABLE 1.1
Dot products of unit vectors in cylindrical and cartesian coordinate systems

	$\mathbf{a}_\rho$	$\mathbf{a}_\phi$	$\mathbf{a}_z$
$\mathbf{a}_x\cdot$	$\cos\phi$	$-\sin\phi$	0
$\mathbf{a}_y\cdot$	$\sin\phi$	$\cos\phi$	0
$\mathbf{a}_z\cdot$	0	0	1

$\mathbf{a}_\rho$ as ϕ, and thus $\mathbf{a}_x \cdot \mathbf{a}_\rho = \cos\phi$, but the angle between $\mathbf{a}_y$ and $\mathbf{a}_\rho$ is $90^\circ - \phi$, and $\mathbf{a}_y \cdot \mathbf{a}_\rho = \cos(90^\circ - \phi) = \sin\phi$. The remaining dot products of the unit vectors are found in a similar manner, and the results are tabulated as functions of ϕ in Table 1.1

Transforming vectors from cartesian to cylindrical coordinates or vice versa is therefore accomplished by using (10) or (11) to change variables, and by using the dot products of the unit vectors given in Table 1.1 to change components. The two steps may be taken in either order.

Example 1.3

Transform the vector $\mathbf{B} = y\mathbf{a}_x - x\mathbf{a}_y + z\mathbf{a}_z$ into cylindrical coordinates.

Solution. The new components are

$$\begin{aligned}
B_\rho &= \mathbf{B}\cdot\mathbf{a}_\rho = y(\mathbf{a}_x\cdot\mathbf{a}_\rho) - x(\mathbf{a}_y\cdot\mathbf{a}_\rho)\\
&= y\cos\phi - x\sin\phi = \rho\sin\phi\cos\phi - \rho\cos\phi\sin\phi = 0\\
B_\phi &= \mathbf{B}\cdot\mathbf{a}_\phi = y(\mathbf{a}_x\cdot\mathbf{a}_\phi) - x(\mathbf{a}_y\cdot\mathbf{a}_\phi)\\
&= -y\sin\phi - x\cos\phi = -\rho\sin^2\phi - \rho\cos^2\phi = -\rho
\end{aligned}$$

Thus,

$$\mathbf{B} = -\rho\mathbf{a}_\phi + z\mathbf{a}_z$$

✔ **D1.5.** (*a*) Give the cartesian coordinates of the point $C(\rho = 4.4, \phi = -115^\circ, z = 2)$. (*b*) Give the cylindrical coordinates of the point $D(x = -3.1, y = 2.6, z = -3)$. (*c*) Specify the distance from C to D.

Ans. $C(x = -1.860, y = -3.99, z = 2)$; $D(\rho = 4.05, \phi = 140.0^\circ, z = -3)$; 8.36

✔ **D1.6.** Transform to cylindrical coordinates: (*a*) $\mathbf{F} = 10\mathbf{a}_x - 8\mathbf{a}_y + 6\mathbf{a}_z$ at point $P(10, -8, 6)$; (*b*) $\mathbf{G} = (2x + y)\mathbf{a}_x - (y - 4x)\mathbf{a}_y$ at point $Q(\rho, \phi, z)$. (*c*) Give the cartesian components of the vector $\mathbf{H} = 20\mathbf{a}_\rho - 10\mathbf{a}_\phi + 3\mathbf{a}_z$ at $P(x = 5, y = 2, z = -1)$.

Ans. $12.81\mathbf{a}_\rho + 6\mathbf{a}_z$; $(2\rho\cos^2\phi - \rho\sin^2\phi + 5\rho\sin\phi\cos\phi)\mathbf{a}_\rho + (4\rho\cos^2\phi - \rho\sin^2\phi - 3\rho\sin\phi\cos\phi)\mathbf{a}_\phi$; $H_x = 22.3$, $H_y = -1.857$, $H_z = 3$

1.9 THE SPHERICAL COORDINATE SYSTEM

We have no two-dimensional coordinate system to help us understand the three-dimensional spherical coordinate system, as we have for the circular cylindrical coordinate system. In certain respects we can draw on our knowledge of the latitude-and-longitude system of locating a place on the surface of the earth, but usually we consider only points on the surface and not those below or above ground.

Let us start by building a spherical coordinate system on the three cartesian axes (Fig. 1.8*a*). We first define the distance from the origin to any point as r. The surface $r =$ constant is a sphere.

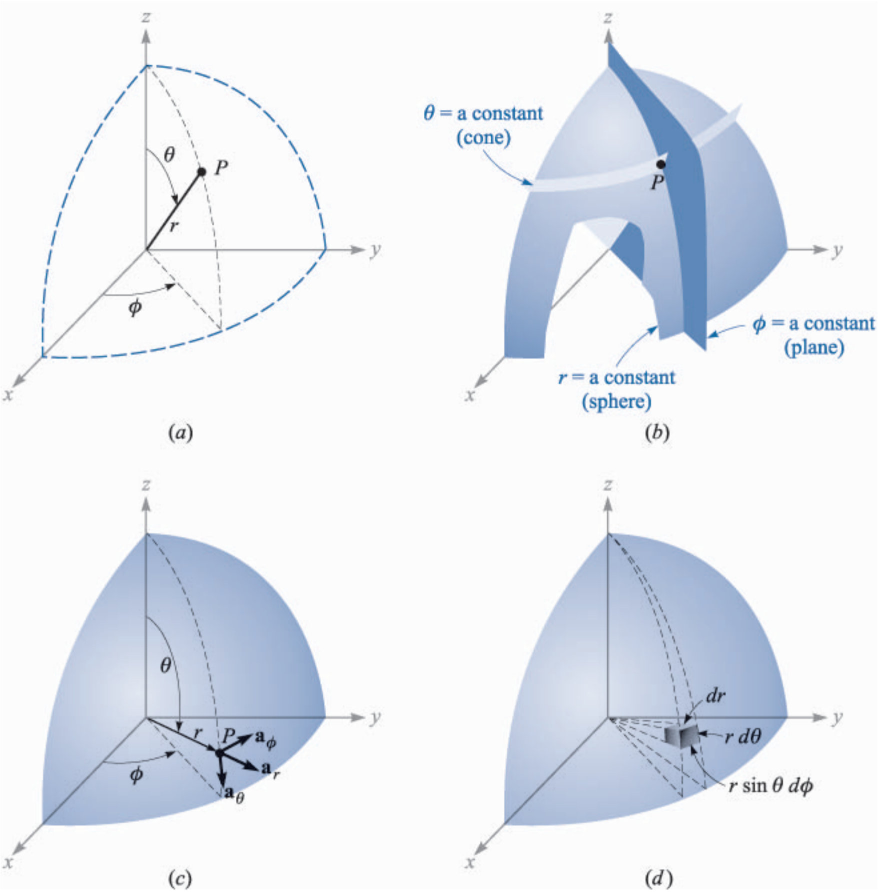


FIGURE 1.8
(*a*) The three spherical coordinates. (*b*) The three mutually perpendicular surfaces of the spherical coordinate system. (*c*) The three unit vectors of spherical coordinates: $\mathbf{a}_r \times \mathbf{a}_\theta = \mathbf{a}_\phi$. (*d*) The differential volume element in the spherical coordinate system.

The second coordinate is an angle θ between the z axis and the line drawn from the origin to the point in question. The surface $\theta =$ constant is a cone, and the two surfaces, cone and sphere, are everywhere perpendicular along their intersection, which is a circle of radius $r \sin \theta$. The coordinate θ corresponds to latitude, except that latitude is measured from the equator and θ is measured from the "North Pole."

The third coordinate ϕ is also an angle and is exactly the same as the angle ϕ of cylindrical coordinates. It is the angle between the x axis and the projection in the $z = 0$ plane of the line drawn from the origin to the point. It corresponds to the angle of longitude, but the angle ϕ increases to the "east." The surface $\phi =$ constant is a plane passing through the $\theta = 0$ line (or the z axis).

We should again consider any point as the intersection of three mutually perpendicular surfaces—a sphere, a cone, and a plane—each oriented in the manner described above. The three surfaces are shown in Fig. 1.8*b*.

Three unit vectors may again be defined at any point. Each unit vector is perpendicular to one of the three mutually perpendicular surfaces and oriented in that direction in which the coordinate increases. The unit vector $\mathbf{a}_r$ is directed radially outward, normal to the sphere $r =$ constant, and lies in the cone $\theta =$ constant and the plane $\phi =$ constant. The unit vector $\mathbf{a}_\theta$ is normal to the conical surface, lies in the plane, and is tangent to the sphere. It is directed along a line of "longitude" and points "south." The third unit vector $\mathbf{a}_\phi$ is the same as in cylindrical coordinates, being normal to the plane and tangent to both the cone and sphere. It is directed to the "east."

The three unit vectors are shown in Fig. 1.8*c*. They are, of course, mutually perpendicular, and a right-handed coordinate system is defined by causing $\mathbf{a}_r \times \mathbf{a}_\theta = \mathbf{a}_\phi$. Our system is right-handed, as an inspection of Fig. 1.8*c* will show, on application of the definition of the cross product. The right-hand rule serves to identify the thumb, forefinger, and middle finger with the direction of increasing r, θ, and ϕ, respectively. (Note that the identification in cylindrical coordinates was with ρ, ϕ, and z, and in cartesian coordinates with x, y, and z). A differential volume element may be constructed in spherical coordinates by increasing r, θ, and ϕ by dr, $d\theta$, and $d\phi$, as shown in Fig. 1.8*d*. The distance between the two spherical surfaces of radius r and $r + dr$ is dr; the distance between the two cones having generating angles of θ and $\theta + d\theta$ is $r d\theta$; and the distance between the two radial planes at angles ϕ and $\phi + d\phi$ is found to be $r \sin \theta d\phi$, after a few moments of trigonometric thought. The surfaces have areas of $r\, dr\, d\theta$, $r \sin \theta\, dr\, d\phi$, and $r^2 \sin \theta\, d\theta\, d\phi$, and the volume is $r^2 \sin \theta\, dr\, d\theta\, d\phi$.

The transformation of scalars from the cartesian to the spherical coordinate system is easily made by using Fig. 1.8*a* to relate the two sets of variables:

$$\begin{aligned} x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \\ z &= r \cos \theta \end{aligned} \tag{15}$$

TABLE 1.2
Dot products of unit vectors in spherical and cartesian coordinate systems

	$\mathbf{a}_r$	$\mathbf{a}_\theta$	$\mathbf{a}_\phi$
$\mathbf{a}_x\cdot$	$\sin\theta\cos\phi$	$\cos\theta\cos\phi$	$-\sin\phi$
$\mathbf{a}_y\cdot$	$\sin\theta\sin\phi$	$\cos\theta\sin\phi$	$\cos\phi$
$\mathbf{a}_z\cdot$	$\cos\theta$	$-\sin\theta$	0

The transformation in the reverse direction is achieved with the help of

$$
\begin{aligned}
r &= \sqrt{x^2+y^2+z^2} && (r \geq 0) \\
\theta &= \cos^{-1}\frac{z}{\sqrt{x^2+y^2+z^2}} && (0^\circ \leq \theta \leq 180^\circ) \\
\phi &= \tan^{-1}\frac{y}{x}
\end{aligned}
\tag{16}
$$

The radius variable r is nonnegative, and θ is restricted to the range from 0° to 180°, inclusive. The angles are placed in the proper quadrants by inspecting the signs of x, y, and z.

The transformation of vectors requires the determination of the products of the unit vectors in cartesian and spherical coordinates. We work out these products from Fig. 1.8*c* and a pinch of trigonometry. Since the dot product of any spherical unit vector with any cartesian unit vector is the component of the spherical vector in the direction of the cartesian vector, the dot products with $\mathbf{a}_z$ are found to be

$$
\begin{aligned}
\mathbf{a}_z \cdot \mathbf{a}_r &= \cos\theta \\
\mathbf{a}_z \cdot \mathbf{a}_\theta &= -\sin\theta \\
\mathbf{a}_z \cdot \mathbf{a}_\phi &= 0
\end{aligned}
$$

The dot products involving $\mathbf{a}_x$ and $\mathbf{a}_y$ require first the projection of the spherical unit vector on the xy plane and then the projection onto the desired axis. For example, $\mathbf{a}_r \cdot \mathbf{a}_x$ is obtained by projecting $\mathbf{a}_r$ onto the xy plane, giving $\sin\theta$, and then projecting $\sin\theta$ on the x axis, which yields $\sin\theta\cos\phi$. The other dot products are found in a like manner, and all are shown in Table 1.2.

Example 1.4

We illustrate this transformation procedure by transforming the vector field $\mathbf{G} = (xz/y)\mathbf{a}_x$ into spherical components and variables.

Solution. We find the three spherical components by dotting $\mathbf{G}$ with the appropriate unit vectors, and we change variables during the procedure:

$$G_r = \mathbf{G} \cdot \mathbf{a}_r = \frac{xz}{y} \mathbf{a}_x \cdot \mathbf{a}_r = \frac{xz}{y} \sin\theta \cos\phi$$

$$= r \sin\theta \cos\theta \frac{\cos^2\phi}{\sin\phi}$$

$$G_\theta = \mathbf{G} \cdot \mathbf{a}_\theta = \frac{xz}{y} \mathbf{a}_x \cdot \mathbf{a}_\theta = \frac{xz}{y} \cos\theta \cos\phi$$

$$= r \cos^2\theta \frac{\cos^2\phi}{\sin\phi}$$

$$G\phi = \mathbf{G} \cdot \mathbf{a}_\phi = \frac{xz}{y} \mathbf{a}_x \cdot \mathbf{a}_\phi = \frac{xz}{y} (-\sin\phi)$$

$$= -r \cos\theta \cos\phi$$

Collecting these results, we have

$$\mathbf{G} = r \cos\theta \cos\phi (\sin\theta \cot\phi \, \mathbf{a}_r + \cos\theta \cot\phi \, \mathbf{a}_\theta - \mathbf{a}_\phi)$$

Appendix A describes the general curvilinear coordinate system of which the cartesian, circular cylindrical, and spherical coordinate systems are special cases. The first section of this appendix could well be scanned now.

✔ **D1.7.** Given the two points, $C(-3, 2, 1)$ and $D(r = 5, \theta = 20^\circ, \phi = -70^\circ)$, find: (*a*) the spherical coordinates of C; (*b*) the cartesian coordinates of D; (*c*) the distance from C to D.

Ans. $C(r = 3.74,\ \theta = 74.5^\circ,\ \phi = 146.3^\circ)$; $D(x = 0.585, y = -1.607, z = 4.70)$; 6.29

✔ **D1.8.** Transform the following vectors to spherical coordinates at the points given: (*a*) $10\mathbf{a}_x$ at $P(x = -3,\ y = 2,\ z = 4)$; (*b*) $10\mathbf{a}_y$ at $Q(\rho = 5, \phi = 30^\circ,\ z = 4)$; (*c*) $10\mathbf{a}_z$ at $M(r = 4, \theta = 110^\circ,\ \phi = 120^\circ)$.

Ans. $-5.57\mathbf{a}_r - 6.18\mathbf{a}_\theta - 5.55\mathbf{a}_\phi$; $3.90\mathbf{a}_r + 3.12\mathbf{a}_\theta + 8.66\mathbf{a}_\phi$; $-3.42\mathbf{a}_r - 9.40\mathbf{a}_\theta$

SUGGESTED REFERENCES

1. Grossman, S. I.: "Calculus," 3d ed., Academic Press and Harcourt Brace Jovanovich, Publishers, Orlando, 1984. Vector algebra and cylindrical and spherical coordinates appear in chap. 17, and vector calculus is introduced in chap. 20.
2. Spiegel, M. R.: "Vector Analysis," Schaum Outline Series, McGraw-Hill Book Company, New York, 1959. A large number of examples and problems with answers are provided in this concise, inexpensive member of an outline series.
3. Swokowski, E. W.: "Calculus with Analytic Geometry," 3d ed., Prindle, Weber, & Schmidt, Boston, 1984. Vector algebra and the cylindrical and spherical coordinate systems are discussed in chap. 14, and vector calculus appears in chap. 18.

4. Thomas, G. B., Jr., and R. L. Finney: "Calculus and Analytic Geometry," 6th ed., Addison-Wesley Publishing Company, Reading, Mass., 1984. Vector algebra and the three coordinate systems we use are discussed in chap. 13. Other vector operations are discussed in chaps. 15 and 17.

PROBLEMS

1.1 Given the vectors $\mathbf{M} = -10\mathbf{a}_x + 4\mathbf{a}_y - 8\mathbf{a}_z$ and $\mathbf{N} = 8\mathbf{a}_x + 7\mathbf{a}_y - 2\mathbf{a}_z$, find: (*a*) a unit vector in the direction of $-\mathbf{M} + 2\mathbf{N}$; (*b*) the magnitude of $5\mathbf{a}_x + \mathbf{N} - 3\mathbf{M}$; (*c*) $|\mathbf{M}||2\mathbf{N}|(\mathbf{M} + \mathbf{N})$.

1.2 Given three points, $A(4, 3, 2)$, $B(-2, 0, 5)$, and $C(7, -2, 1)$: (*a*) specify the vector $\mathbf{A}$ extending from the origin to point A; (*b*) give a unit vector extending from the origin toward the midpoint of line AB; (*c*) calculate the length of the perimeter of triangle ABC.

1.3 The vector from the origin to point A is given as $6\mathbf{a}_x - 2\mathbf{a}_y - 4\mathbf{a}_z$, and the unit vector directed from the origin toward point B is $\left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)$. If points A and B are 10 units apart, find the coordinates of point B.

1.4 Given points $A(8, -5, 4)$ and $B(-2, 3, 2)$, find: (*a*) the distance from A to B; (*b*) a unit vector directed from A towards B; (*c*) a unit vector directed from the origin toward the midpoint of the line AB; (*d*) the coordinates of the point on the line connecting A to B at which the line intersects the plane $z = 3$.

1.5 A vector field is specified as $\mathbf{G} = 24xy\mathbf{a}_x + 12(x^2 + 2)\mathbf{a}_y + 18z^2\mathbf{a}_z$. Given two points, $P(1, 2, -1)$ and $Q(-2, 1, 3)$, find: (*a*) $\mathbf{G}$ at P; (*b*) a unit vector in the direction of $\mathbf{G}$ at Q; (*c*) a unit vector directed from Q toward P; (*d*) the equation of the surface on which $|\mathbf{G}| = 60$.

1.6 For the $\mathbf{G}$ field given in Prob. 1.5 above, make sketches of G_x G_y, G_z and $|\mathbf{G}|$ along the line $y = 1$, $z = 1$, for $0 \le x \le 2$.

1.7 Given the vector field $\mathbf{E} = 4zy^2 \cos 2x\mathbf{a}_x + 2zy \sin 2x\mathbf{a}_y + y^2 \sin 2x\mathbf{a}_z$, find, for the region $|x|$, $|y|$, and $|z| < 2$: (*a*) the surfaces on which $E_y = 0$; (*b*) the region in which $E_y = E_z$; (*c*) the region for which $\mathbf{E} = 0$.

1.8 Two vector fields are $\mathbf{F} = -10\mathbf{a}_x + 20x(y-1)\mathbf{a}_y$ and $\mathbf{G} = 2x^2y\mathbf{a}_x - 4\mathbf{a}_y + z\mathbf{a}_z$. For the point $P(2, 3, -4)$, find: (*a*) $|\mathbf{F}|$; (*b*) $|\mathbf{G}|$; (*c*) a unit vector in the direction of $\mathbf{F} - \mathbf{G}$; (*d*) a unit vector in the direction of $\mathbf{F} + \mathbf{G}$.

1.9 A field is given as $\mathbf{G} = \dfrac{25}{x^2 + y^2}(x\mathbf{a}_x + y\mathbf{a}_y)$. Find: (*a*) a unit vector in the direction of $\mathbf{G}$ at $P(3, 4, -2)$; (*b*) the angle between $\mathbf{G}$ and $\mathbf{a}_x$ at P; (*c*) the value of the double integral $\int_{x=0}^{4} \int_{z=0}^{2} \mathbf{G} \cdot dx\,dz\,\mathbf{a}_y$ on the plane $y = 7$.

1.10 Use the definition of the dot product to find the interior angles at A and B of the triangle defined by the three points: $A(1, 3, -2)$, $B(-2, 4, 5)$, and $C(0, -2, 1)$.

1.11 Given the points $M(0.1, -0.2, -0.1)$, $N(-0.2, 0.1, 0.3)$, and $P(0.4, 0, 0.1)$, find: (*a*) the vector $\mathbf{R}_{MN}$; (*b*) the dot product $\mathbf{R}_{MN} \cdot \mathbf{R}_{MP}$; (*c*) the scalar projection of $\mathbf{R}_{MN}$ on $\mathbf{R}_{MP}$; (*d*) the angle between $\mathbf{R}_{MN}$ and $\mathbf{R}_{MP}$.

1.12 Given points $A(10, 12, -6)$, $B(16, 8, -2)$, $C(8, 1, 4)$, and $D(-2, -5, 8)$, determine: (*a*) the vector projection of $\mathbf{R}_{AB} + \mathbf{R}_{BC}$ on $\mathbf{R}_{AD}$; (*b*) the vector projection of $\mathbf{R}_{AB} + \mathbf{R}_{BC}$ on $\mathbf{R}_{DC}$; (*c*) the angle between $\mathbf{R}_{DA}$ and $\mathbf{R}_{DC}$.

1.13 (*a*) Find the vector component of $\mathbf{F} = 10\mathbf{a}_x - 6\mathbf{a}_y + 5\mathbf{a}_z$ that is parallel to $\mathbf{G} = 0.1\mathbf{a}_x + 0.2\mathbf{a}_y + 0.3\mathbf{a}_z$. (*b*) Find the vector component of $\mathbf{F}$ that is perpendicular to $\mathbf{G}$. (*c*) Find the vector component of $\mathbf{G}$ that is perpendicular to $\mathbf{F}$.

1.14 The three vertices of a regular tetrahedron are located at $O(0, 0, 0)$, $A(0, 1, 0)$, $B(0.5\sqrt{3}, 0.5, 0)$, and $C(\sqrt{3}/6, 0.5, \sqrt{2/3})$. (*a*) Find a unit vector perpendicular (outward) to face ABC; (*b*) Find the area of face ABC.

1.15 Three vectors extending from the origin are given as $\mathbf{r}_1 = 7\mathbf{a}_x + 3\mathbf{a}_y - 2\mathbf{a}_z$, $\mathbf{r}_2 = -2\mathbf{a}_x + 7\mathbf{a}_y - 3\mathbf{a}_z$, and $\mathbf{r}_3 = 2\mathbf{a}_x - 2\mathbf{a}_y + 3\mathbf{a}_z$. Find: (*a*) a unit vector perpendicular to both $\mathbf{r}_1$ and $\mathbf{r}_2$; (*b*) a unit vector perpendicular to the vectors $\mathbf{r}_1 - \mathbf{r}_2$ and $\mathbf{r}_2 - \mathbf{r}_3$; (*c*) the area of the triangle defined by $\mathbf{r}_1$ and $\mathbf{r}_2$; (*d*) the area of the triangle defined by the heads of $\mathbf{r}_1$, $\mathbf{r}_2$, and $\mathbf{r}_3$.

1.16 Describe the surface defined by the equation: (*a*) $\mathbf{r} \cdot \mathbf{a}_x = 2$, where $\mathbf{r} = x\mathbf{a}_x + y\mathbf{a}_y + z\mathbf{a}_z$; (*b*) $|\mathbf{r} \times \mathbf{a}_x| = 2$.

1.17 Point $A(-4, 2, 5)$ and the two vectors, $\mathbf{R}_{AM} = 20\mathbf{a}_x + 18\mathbf{a}_y - 10\mathbf{a}_z$ and $\mathbf{R}_{AN} = -10\mathbf{a}_x + 8\mathbf{a}_y + 15\mathbf{a}_z$, define a triangle. (*a*) Find a unit vector perpendicular to the triangle. (*b*) Find a unit vector in the plane of the triangle and perpendicular to $\mathbf{R}_{AN}$. (*c*) Find a unit vector in the plane of the triangle that bisects the interior angle at A.

1.18 Given points $A(\rho = 5, \phi = 70^\circ, z = -3)$ and $B(\rho = 2, \phi = -30^\circ, z = 1)$, find: (*a*) a unit vector in cartesian coordinates at A directed toward B; (*b*) a unit vector in cylindrical coordinates at A directed toward B; (*c*) a unit vector in cylindrical coordinates at B directed toward A.

1.19 (*a*) Express the vector field $\mathbf{D} = (x^2 + y^2)^{-1}(x\mathbf{a}_x + y\mathbf{a}_y)$ in cylindrical components and cylindrical variables. (*b*) Evaluate $\mathbf{D}$ at the point where $\rho = 2$, $\varphi = 0.2\pi$ (rad), and $z = 5$. Express the result in both cylindrical and cartesian components.

1.20 Express in cartesian components: (*a*) the vector at $A(\rho = 4, \phi = 40^\circ, z = -2)$ that extends to $B(\rho = 5, \phi = -110^\circ, z = 2)$; (*b*) a unit vector at B directed toward A; (*c*) a unit vector at B directed toward the origin.

1.21 Express in cylindrical components: (*a*) the vector from $C(3, 2, -7)$ to $D(-1, -4, 2)$; (*b*) a unit vector at D directed toward C; (*c*) a unit vector at D directed toward the origin.

1.22 A field is given in cylindrical coordinates as $\mathbf{F} = \left[\dfrac{40}{\rho^2 + 1} + 3(\cos\phi + \sin\phi)\right]\mathbf{a}_\rho + 3(\cos\phi - \sin\phi)\mathbf{a}_\phi - 2\mathbf{a}_z$. Prepare simple sketches of $|\mathbf{F}|$: (*a*) vs ϕ with $\rho = 3$; (*b*) vs ρ with $\phi = 0$; (*c*) vs ρ with $\phi = 45^\circ$.

1.23 The surfaces $\rho = 3$ and 5, $\phi = 100^\circ$ and 130°, and $z = 3$ and 4.5 identify a closed surface. (*a*) Find the volume enclosed. (*b*) Find the total area of the enclosing surface. (*c*) Find the total length of the twelve edges of the

surface. (*d*) Find the length of the longest straight line that lies entirely within the volume.

1.24 At point $P(-3, -4, 5)$, express that vector that extends from P to $Q(2, 0, -1)$ in: (*a*) rectangular coordinates; (*b*) cylindrical coordinates; (*c*) spherical coordinates. (*d*) Show that each of these vectors has the same magnitude.

1.25 Let $\mathbf{E} = \frac{1}{r^2}\left(\cos\phi\,\mathbf{a}_r + \frac{\sin\phi}{\sin\theta}\mathbf{a}_\phi\right)$. Given point $P(r = 0.8,\ \theta = 30^\circ,\ \phi = 45^\circ)$, determine: (*a*) $\mathbf{E}$ at P; (*b*) $|\mathbf{E}|$ at P; (*c*) a unit vector in the direction of $\mathbf{E}$ at P.

1.26 (*a*) Determine an expression for $\mathbf{a}_y$ in spherical coordinates at $P(r = 4, \theta = 0.2\pi, \phi = 0.8\pi)$. (*b*) Express $\mathbf{a}_r$ in cartesian components at P.

1.27 The surfaces $r = 2$ and 4, $\theta = 30^\circ$ and 50°, and $\phi = 20^\circ$ and 60° identify a closed surface. (*a*) Find the enclosed volume. (*b*) Find the total area of the enclosing surface. (*c*) Find the total length of the twelve edges of the surface. (*d*) Find the length of the longest straight line that lies entirely within the volume.

1.28 (*a*) Determine the cartesian components of the vector from $A(r = 5, \theta = 110^\circ, \phi = 200^\circ)$ to $B(r = 7, \theta = 30^\circ, \phi = 70^\circ)$. (*b*) Find the spherical components of the vector at $P(2, -3, 4)$ extending to $Q(-3, 2, 5)$. (*c*) If $\mathbf{D} = 5\mathbf{a}_r - 3\mathbf{a}_\theta + 4\mathbf{a}_\phi$, find $\mathbf{D}\cdot\mathbf{a}_\rho$ at $M(1, 2, 3)$.

1.29 Express the unit vector $\mathbf{a}_x$ in spherical components at the point: (*a*) $r = 2$, $\theta = 1$ rad, $\phi = 0.8$ rad; (*b*) $x = 3$, $y = 2$, $z = -1$; (*c*) $\rho = 2.5$, $\phi = 0.7$ rad, $z = 1.5$.

1.30 Given $A(r = 20, \theta = 30^\circ, \phi = 45^\circ)$ and $B(r = 30, \theta = 115^\circ, \phi = 160^\circ)$, find: (*a*) $|\mathbf{R}_{AB}|$; (*b*) $|\mathbf{R}_{AC}|$, given $C(r = 20, \theta = 90^\circ, \phi = 45^\circ)$; (*c*) the distance from A to C on a great circle path.

CHAPTER

2

COULOMB'S LAW AND ELECTRIC FIELD INTENSITY

Now that we have formulated a new language in the first chapter, we shall establish a few basic principles of electricity and attempt to describe them in terms of it. If we had used vector calculus for several years and already had a few correct ideas about electricity and magnetism, we might jump in now with both feet and present a handful of equations, including Maxwell's equations and a few other auxiliary equations, and proceed to describe them physically by virtue of our knowledge of vector analysis. This is perhaps the ideal way, starting with the most general results and then showing that Ohm's, Gauss's, Coulomb's, Faraday's, Ampère's, Biot-Savart's, Kirchhoff's, and a few less familiar laws are all special cases of these equations. It is philosophically satisfying to have the most general result and to feel that we are able to obtain the results for any special case at will. However, such a jump would lead to many frantic cries of "Help" and not a few drowned students.

Instead we shall present at decent intervals the experimental laws mentioned above, expressing each in vector notation, and use these laws to solve a

number of simple problems. In this way our familiarity with both vector analysis and electric and magnetic fields will gradually increase, and by the time we have finally reached our handful of general equations, little additional explanation will be required. The entire field of electromagnetic theory is then open to us, and we may use Maxwell's equations to describe wave propagation, radiation from antennas, skin effect, waveguides and transmission lines, and travelling-wave tubes, and even to obtain a new insight into the ordinary power transformer.

In this chapter we shall restrict our attention to *static* electric fields in *vacuum* or *free space*. Such fields, for example, are found in the focusing and deflection systems of electrostatic cathode-ray tubes. For all practical purposes, our results will also be applicable to air and other gases. Other materials will be introduced in Chap. 5, and time-varying fields will be introduced in Chap. 10.

We shall begin by describing a quantitative experiment performed in the seventeenth century.

2.1 THE EXPERIMENTAL LAW OF COULOMB

Records from at least 600 B.C. show evidence of the knowledge of static electricity. The Greeks were responsible for the term "electricity," derived from their word for amber, and they spent many leisure hours rubbing a small piece of amber on their sleeves and observing how it would then attract pieces of fluff and stuff. However, their main interest lay in philosophy and logic, not in experimental science, and it was many centuries before the attracting effect was considered to be anything other than magic or a "life force."

Dr. Gilbert, physician to Her Majesty the Queen of England, was the first to do any true experimental work with this effect and in 1600 stated that glass, sulfur, amber, and other materials which he named would "not only draw to themselves straws and chaff, but all metals, wood, leaves, stone, earths, even water and oil."

Shortly thereafter a colonel in the French Army Engineers, Colonel Charles Coulomb, a precise and orderly minded officer, performed an elaborate series of experiments using a delicate torsion balance, invented by himself, to determine quantitatively the force exerted between two objects, each having a static charge of electricity. His published result is now known to many high school students and bears a great similarity to Newton's gravitational law (discovered about a hundred years earlier). Coulomb stated that the force between two very small objects separated in a vacuum or free space by a distance which is large compared to their size is proportional to the charge on each and inversely proportional to the square of the distance between them, or

$$F = k\frac{Q_1 Q_2}{R^2}$$

where Q_1 and Q_2 are the positive or negative quantities of charge, R is the separation, and k is a proportionality constant. If the International System of

Units[1] (SI) is used, Q is measured in coulombs (C), R is in meters (m), and the force should be newtons (N). This will be achieved if the constant of proportionality k is written as

$$k = \frac{1}{4\pi\epsilon_0}$$

The factor 4π will appear in the denominator of Coulomb's law but will not appear in the more useful equations (including Maxwell's equations) which we shall obtain with the help of Coulomb's law. The new constant ϵ_0 is called the *permittivity of free space* and has the magnitude, measured in farads per meter (F/m),

$$\epsilon_0 = 8.854 \times 10^{-12} \doteq \frac{1}{36\pi} 10^{-9} \qquad \text{F/m} \tag{1}$$

The quantity ϵ_0 is not dimensionless, for Coulomb's law shows that it has the label $\text{C}^2/\text{N}\cdot\text{m}^2$. We shall later define the farad and show that it has the dimensions $\text{C}^2/\text{N}\cdot\text{m}$; we have anticipated this definition by using the unit F/m in (1) above.

Coulomb's law is now

$$F = \frac{Q_1 Q_2}{4\pi\epsilon_0 R^2} \tag{2}$$

Not all SI units are as familiar as the English units we use daily, but they are now standard in electrical engineering and physics. The newton is a unit of force that is equal to 0.2248 lb_f, and is the force required to give a 1-kilogram (kg) mass an acceleration of 1 meter per second per second (m/s^2). The coulomb is an extremely large unit of charge, for the smallest known quantity of charge is that of the electron (negative) or proton (positive), given in mks units as 1.602×10^{-19} C; hence a negative charge of one coulomb represents about 6×10^{18} electrons.[2] Coulomb's law shows that the force between two charges of one coulomb each, separated by one meter, is 9×10^9 N, or about one million tons. The electron has a rest mass of 9.109×10^{-31} kg and has a radius of the order of magnitude of 3.8×10^{-15} m. This does not mean that the electron is spherical in shape, but merely serves to describe the size of the region in which a slowly moving electron has the greatest probability of being found. All other

[1] The International System of Units (an mks system) is described in Appendix B. Abbreviations for the units are given in Table B.1. Conversions to other systems of units are given in Table B.2, while the prefixes designating powers of ten in S1 appear in Table B.3.

[2] The charge and mass of an electron and other physical constants are tabulated in Table C.4 of Appendix C.

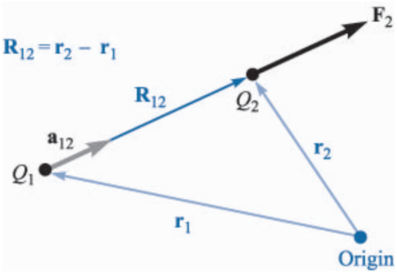


FIGURE 2.1
If Q_1 and Q_2 have like signs, the vector force $\mathbf{F}_2$ on Q_2 is in the same direction as the vector $\mathbf{R}_{12}$.

known charged particles, including the proton, have larger masses, and larger radii, and occupy a probabilistic volume larger than does the electron.

In order to write the vector form of (2), we need the additional fact (furnished also by Colonel Coulomb) that the force acts along the line joining the two charges and is repulsive if the charges are alike in sign and attractive if they are of opposite sign. Let the vector $\mathbf{r}_1$ locate Q_1 while $\mathbf{r}_2$ locates Q_2. Then the vector $\mathbf{R}_{12} = \mathbf{r}_2 - \mathbf{r}_1$ represents the directed line segment from Q_1 to Q_2, as shown in Fig. 2.1. The vector $\mathbf{F}_2$ is the force on Q_2 and is shown for the case where Q_1 and Q_2 have the same sign. The vector form of Coulomb's law is

$$\mathbf{F}_2 = \frac{Q_1 Q_2}{4\pi\epsilon_0 R_{12}^2}\mathbf{a}_{12} \tag{3}$$

where $\mathbf{a}_{12}$ = a unit vector in the direction of R_{12}, or

$$\mathbf{a}_{12} = \frac{\mathbf{R}_{12}}{|\mathbf{R}_{12}|} = \frac{\mathbf{R}_{12}}{R_{12}} = \frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_2 - \mathbf{r}_1|} \tag{4}$$

Example 2.1

Let us illustrate the use of the vector form of Coulomb's law by locating a charge of $Q_1 = 3 \times 10^{-4}$ C at $M(1, 2, 3)$ and a charge of $Q_2 = -10^{-4}$ C at $N(2, 0, 5)$ in a vacuum. We desire the force exerted on Q_2 by Q_1.

Solution. We shall make use of (3) and (4) to obtain the vector force. The vector $\mathbf{R}_{12}$ is

$$\mathbf{R}_{12} = \mathbf{r}_2 - \mathbf{r}_1 = (2-1)\mathbf{a}_x + (0-2)\mathbf{a}_y + (5-3)\mathbf{a}_z = \mathbf{a}_x - 2\mathbf{a}_y + 2\mathbf{a}_z$$

leading to $|\mathbf{R}_{12}| = 3$, and the unit vector, $\mathbf{a}_{12} = \frac{1}{3}(\mathbf{a}_x - 2\mathbf{a}_y + 2\mathbf{a}_z)$. Thus,

$$\mathbf{F}_2 = \frac{3 \times 10^{-4}(-10^{-4})}{4\pi(1/36\pi)10^{-9} \times 3^2}\left(\frac{\mathbf{a}_x - 2\mathbf{a}_y + 2\mathbf{a}_z}{3}\right)$$
$$= -30\left(\frac{\mathbf{a}_x - 2\mathbf{a}_y + 2\mathbf{a}_z}{3}\right) \text{ N}$$

The magnitude of the force is 30 N (or about 7 lb_f), and the direction is specified by the unit vector, which has been left in parentheses to display the magnitude of the force. The force on Q_2 may also be considered as three component forces,

$$\mathbf{F}_2 = -10\mathbf{a}_x + 20\mathbf{a}_y - 20\mathbf{a}_z$$

The force expressed by Coulomb's law is a mutual force, for each of the two charges experiences a force of the same magnitude, although of opposite direction. We might equally well have written

$$\mathbf{F}_1 = -\mathbf{F}_2 = \frac{Q_1 Q_2}{4\pi\epsilon_0 R_{12}^2}\mathbf{a}_{21} = -\frac{Q_1 Q_2}{4\pi\epsilon_0 R_{12}^2}\mathbf{a}_{12} \tag{5}$$

Coulomb's law is linear, for if we multiply Q_1 by a factor n, the force on Q_2 is also multiplied by the same factor n. It is also true that the force on a charge in the presence of several other charges is the sum of the forces on that charge due to each of the other charges acting alone.

✔ **D2.1.** A charge $Q_A = -20\,\mu\text{C}$ is located at $A(-6, 4, 7)$, and a charge $Q_B = 50\,\mu\text{C}$ is at $B(5, 8, -2)$ in free space. If distances are given in meters, find: (*a*) $\mathbf{R}_{AB}$; (*b*) R_{AB}. Determine the vector force exerted on Q_A by Q_B if $\epsilon_0 =$: (*c*) $10^{-9}/(36\pi)$ F/m; (*d*) 8.854×10^{-12} F/m.

Ans. $11\mathbf{a}_x + 4\mathbf{a}_y - 9\mathbf{a}_z$ m; 14.76 m; $30.76\mathbf{a}_x + 11.184\mathbf{a}_y - 25.16\mathbf{a}_z$ mN; $30.72\mathbf{a}_x + 11.169\mathbf{a}_y - 25.13\mathbf{a}_z$ mN

2.2 ELECTRIC FIELD INTENSITY

If we now consider one charge fixed in position, say Q_1, and move a second charge slowly around, we note that there exists everywhere a force on this second charge; in other words, this second charge is displaying the existence of a force *field*. Call this second charge a test charge Q_t. The force on it is given by Coulomb's law,

$$\mathbf{F}_t = \frac{Q_1 Q_t}{4\pi\epsilon_0 R_{1t}^2}\mathbf{a}_{1t}$$

Writing this force as a force per unit charge gives

$$\frac{\mathbf{F}_t}{Q_t} = \frac{Q_1}{4\pi\epsilon_0 R_{1t}^2}\mathbf{a}_{1t} \tag{6}$$

The quantity on the right side of (6) is a function only of Q_1 and the directed line segment from Q_1 to the position of the test charge. This describes a vector field and is called the *electric field intensity*.

We define the electric field intensity as the vector force on a unit positive test charge. We would not *measure* it experimentally by finding the force on a 1-C test charge, however, for this would probably cause such a force on Q_1 as to change the position of that charge.

Electric field intensity must be measured by the unit newtons per coulomb—the force per unit charge. Again anticipating a new dimensional quantity, the *volt* (V), to be presented in Chap. 4 and having the label of joules per coulomb (J/C) or newton-meters per coulomb (N·m/C), we shall at once measure electric field intensity in the practical units of volts per meter (V/m). Using a capital letter **E** for electric field intensity, we have finally

$$\boxed{\mathbf{E} = \frac{\mathbf{F}_t}{Q_t}} \tag{7}$$

$$\mathbf{E} = \frac{Q_1}{4\pi\epsilon_0 R_{1t}^2}\mathbf{a}_{1t} \tag{8}$$

Equation (7) is the defining expression for electric field intensity, and (8) is the expression for the electric field intensity due to a single point charge Q_1 in a vacuum. In the succeeding sections we shall obtain and interpret expressions for the electric field intensity due to more complicated arrangements of charge, but now let us see what information we can obtain from (8), the field of a single point charge.

First, let us dispense with most of the subscripts in (8), reserving the right to use them again any time there is a possibility of misunderstanding:

$$\boxed{\mathbf{E} = \frac{Q}{4\pi\epsilon_0 R^2}\mathbf{a}_R} \tag{9}$$

We should remember that R is the magnitude of the vector **R**, the directed line segment from the point at which the point charge Q is located to the point at which **E** is desired, and $\mathbf{a}_R$ is a unit vector in the **R** direction.[3]

Let us arbitrarily locate Q_1 at the center of a spherical coordinate system. The unit vector $\mathbf{a}_R$ then becomes the radial unit vector $\mathbf{a}_r$, and R is r. Hence

$$\mathbf{E} = \frac{Q_1}{4\pi\epsilon_0 r^2}\mathbf{a}_r \tag{10}$$

or

$$E_r = \frac{Q_1}{4\pi\epsilon_0 r^2}$$

The field has a single radial component, and its inverse-square-law relationship is quite obvious.

[3] We firmly intend to avoid confusing r and $\mathbf{a}_r$ with R and $\mathbf{a}_R$. The first two refer specifically to the spherical coordinate system, whereas R and $\mathbf{a}_R$ do not refer to any coordinate system—the choice is still available to us.

Writing these expressions in cartesian coordinates for a charge Q at the origin, we have $\mathbf{R} = \mathbf{r} = x\mathbf{a}_x + y\mathbf{a}_y + z\mathbf{a}_z$ and $\mathbf{a}_R = \mathbf{a}_r = (x\mathbf{a}_x + y\mathbf{a}_y + z\mathbf{a}_z)/\sqrt{x^2 + y^2 + z^2}$; therefore,

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0(x^2 + y^2 + z^2)}\left(\frac{x}{\sqrt{x^2 + y^2 + z^2}}\mathbf{a}_x + \frac{y}{\sqrt{x^2 + y^2 + z^2}}\mathbf{a}_y + \frac{z}{\sqrt{x^2 + y^2 + z^2}}\mathbf{a}_z\right) \tag{11}$$

This expression no longer shows immediately the simple nature of the field, and its complexity is the price we pay for solving a problem having spherical symmetry in a coordinate system with which we may (temporarily) have more familiarity.

Without using vector analysis, the information contained in (11) would have to be expressed in three equations, one for each component, and in order to obtain the equation we would have to break up the magnitude of the electric field intensity into the three components by finding the projection on each coordinate axis. Using vector notation, this is done automatically when we write the unit vector.

If we consider a charge which is *not* at the origin of our coordinate system, the field no longer possesses spherical symmetry (nor cylindrical symmetry, unless the charge lies on the z axis), and we might as well use cartesian coordinates. For a charge Q located at the source point $\mathbf{r}' = x'\mathbf{a}_x + y'\mathbf{a}_y + z'\mathbf{a}_z$, as illustrated in Fig. 2.2, we find the field at a general field point $\mathbf{r} = x\mathbf{a}_x + y\mathbf{a}_y + z\mathbf{a}_z$ by expressing $\mathbf{R}$ as $\mathbf{r} - \mathbf{r}'$, and then

$$\mathbf{E}(\mathbf{r}) = \frac{Q}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|^2}\frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} = \frac{Q(\mathbf{r} - \mathbf{r}')}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|^3}$$
$$= \frac{Q[(x - x')\mathbf{a}_x + (y - y')\mathbf{a}_y + (z - z')\mathbf{a}_z]}{4\pi\epsilon_0[(x - x')^2 + (y - y')^2 + (z - z')^2]^{3/2}} \tag{12}$$

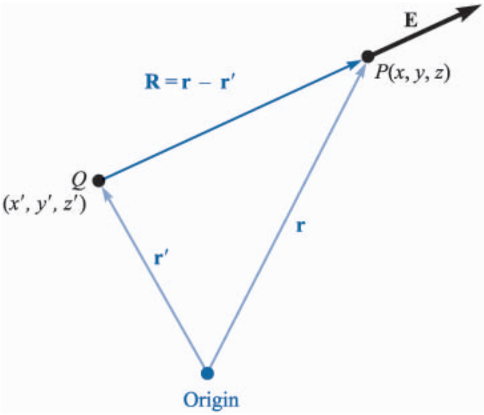


FIGURE 2.2
The vector $\mathbf{r}'$ locates the point charge Q, the vector $\mathbf{r}$ identifies the general point in space $P(x, y, z)$, and the vector $\mathbf{R}$ from Q to $P(x, y, z)$ is then $\mathbf{R} = \mathbf{r} - \mathbf{r}'$.

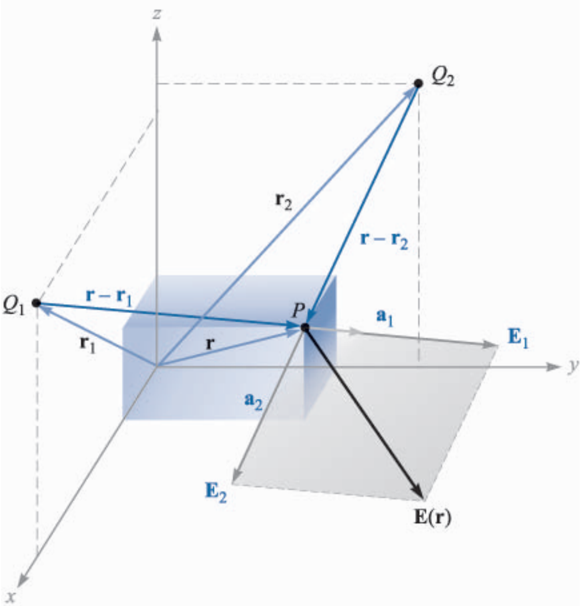


FIGURE 2.3
The vector addition of the total electric field intensity at P due to Q_1 and Q_2 is made possible by the linearity of Coulomb's law.

Earlier, we defined a vector field as a vector function of a position vector, and this is emphasized by letting $\mathbf{E}$ be symbolized in functional notation by $\mathbf{E}(\mathbf{r})$.

Equation (11) is merely a special case of (12), where $x' = y' = z' = 0$.

Since the coulomb forces are linear, the electric field intensity due to two point charges, Q_1 at $\mathbf{r}_1$ and Q_2 at $\mathbf{r}_2$, is the sum of the forces on Q_t caused by Q_1 and Q_2 acting alone, or

$$\mathbf{E}(\mathbf{r}) = \frac{Q_1}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_1|^2}\mathbf{a}_1 + \frac{Q_2}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_2|^2}\mathbf{a}_2$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are unit vectors in the direction of $(\mathbf{r}-\mathbf{r}_1)$ and $(\mathbf{r}-\mathbf{r}_2)$, respectively. The vectors $\mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}-\mathbf{r}_1, \mathbf{r}-\mathbf{r}_2, \mathbf{a}_1$, and $\mathbf{a}_2$ are shown in Fig. 2.3.

If we add more charges at other positions, the field due to n point charges is

$$\mathbf{E}(\mathbf{r}) = \frac{Q_1}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_1|^2}\mathbf{a}_1 + \frac{Q_2}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_2|^2}\mathbf{a}_2 + \ldots + \frac{Q_n}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_n|^2}\mathbf{a}_n \qquad (13)$$

This expression takes up less space when we use a summation sign $\sum$ and a summing integer m which takes on all integral values between 1 and n,

$$\mathbf{E}(\mathbf{r}) = \sum_{m=1}^{n} \frac{Q_m}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_m|^2}\mathbf{a}_m \qquad (14)$$

When expanded, (14) is identical with (13), and students unfamiliar with summation signs should check that result.

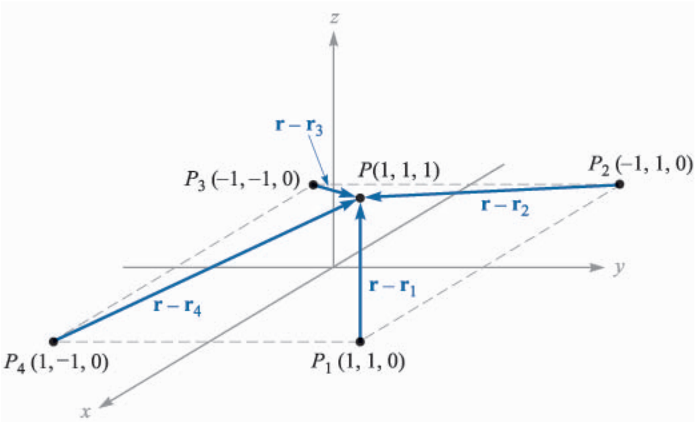


FIGURE 2.4
A symmetrical distribution of four identical 3-nC point charges produces a field at P, $\mathbf{E} = 6.82\mathbf{a}_x + 6.82\mathbf{a}_y + 32.8\mathbf{a}_z$ V/m.

Example 2.2

In order to illustrate the application of (13) or (14), let us find $\mathbf{E}$ at $P(1, 1, 1)$ caused by four identical 3-nC (nanocoulomb) charges located at $P_1(1, 1, 0)$, $P_2(-1, 1, 0)$, $P_3(-1, -1, 0)$, and $P_4(1, -1, 0)$, as shown in Fig. 2.4.

Solution. We find that $\mathbf{r} = \mathbf{a}_x + \mathbf{a}_y + \mathbf{a}_z$, $\mathbf{r}_1 = \mathbf{a}_x + \mathbf{a}_y$, and thus $\mathbf{r} - \mathbf{r}_1 = \mathbf{a}_z$. The magnitudes are: $|\mathbf{r} - \mathbf{r}_1| = 1$, $|\mathbf{r} - \mathbf{r}_2| = \sqrt{5}$, $|\mathbf{r} - \mathbf{r}_3| = 3$, and $|\mathbf{r} - \mathbf{r}_4| = \sqrt{5}$. Since $Q/4\pi\epsilon_0 = 3 \times 10^{-9}/(4\pi \times 8.854 \times 10^{-12}) = 26.96\ \text{V} \cdot \text{m}$, we may now use (13) or (14) to obtain

$$\mathbf{E} = 26.96\left[\frac{\mathbf{a}_z}{1}\frac{1}{1^2} + \frac{2\mathbf{a}_x + \mathbf{a}_z}{\sqrt{5}}\frac{1}{(\sqrt{5})^2} + \frac{2\mathbf{a}_x + 2\mathbf{a}_y + \mathbf{a}_z}{3}\frac{1}{3^2} + \frac{2\mathbf{a}_y + \mathbf{a}_z}{\sqrt{5}}\frac{1}{(\sqrt{5})^2}\right]$$

or

$$\mathbf{E} = 6.82\mathbf{a}_x + 6.82\mathbf{a}_y + 32.8\mathbf{a}_z \ \text{V/m}$$

✔ **D2.2.** A charge of $-0.3\ \mu$C is located at $A(25, -30, 15)$ (in cm), and a second charge of $0.5\ \mu$C is at $B(-10, 8, 12)$ cm. Find $\mathbf{E}$ at: (*a*) the origin; (*b*) $P(15, 20, 50)$ cm.

Ans. $92.3\mathbf{a}_x - 77.6\mathbf{a}_y - 105.3\mathbf{a}_z$ kV/m; $32.9\mathbf{a}_x + 5.94\mathbf{a}_y + 19.69\mathbf{a}_z$ kV/m

✔ **D2.3.** Evaluate the sums: (*a*) $\sum_{m=0}^{5} \frac{1 + (-1)^m}{m^2 + 1}$; (*b*) $\sum_{m=1}^{4} \frac{(0.1)^m + 1}{(4 + m^2)^{1.5}}$

Ans. 2.52; 0.1948

2.3 FIELD DUE TO A CONTINUOUS VOLUME CHARGE DISTRIBUTION

If we now visualize a region of space filled with a tremendous number of charges separated by minute distances, such as the space between the control grid and the cathode in the electron-gun assembly of a cathode-ray tube operating with space charge, we see that we can replace this distribution of very small particles with a smooth continuous distribution described by a *volume charge density*, just as we describe water as having a density of 1 g/cm^3 (gram per cubic centimeter) even though it consists of atomic- and molecular-sized particles. We are able to do this only if we are uninterested in the small irregularities (or ripples) in the field as we move from electron to electron or if we care little that the mass of the water actually increases in small but finite steps as each new molecule is added.

This is really no limitation at all, because the end results for electrical engineers are almost always in terms of a current in a receiving antenna, a voltage in an electronic circuit, or a charge on a capacitor, or in general in terms of some large-scale *macroscopic* phenomenon. It is very seldom that we must know a current electron by electron.[4]

We denote volume charge density by ρ_v, having the units of coulombs per cubic meter (C/m^3).

The small amount of charge ΔQ in a small volume Δv is

$$\Delta Q = \rho_v \Delta v \tag{15}$$

and we may define ρ_v mathematically by using a limiting process on (15),

$$\boxed{\rho_v = \lim_{\Delta v \to 0} \frac{\Delta Q}{\Delta v}} \tag{16}$$

The total charge within some finite volume is obtained by integrating throughout that volume,

$$\boxed{Q = \int_{\text{vol}} \rho_v \, dv} \tag{17}$$

Only one integral sign is customarily indicated, but the differential *dv* signifies integration throughout a volume, and hence a triple integration. Fortunately, we may be content for the most part with no more than the indicated integration, for multiple integrals are very difficult to evaluate in all but the most symmetrical problems.

---

[4] A study of the noise generated by electrons or ions in transistors, vacuum tubes, and resistors, however, requires just such an examination of the charge.

Example 2.3

As an example of the evaluation of a volume integral, we shall find the total charge contained in a 2-cm length of the electron beam shown in Fig. 2.5.

Solution. From the illustration, we see that the charge density is

$$\rho_v = -5 \times 10^{-6} e^{-10^5 \rho z} \text{ C/m}^2$$

The volume differential in cylindrical coordinates is given in Sec. 1.8; therefore,

$$Q = \int_{0.02}^{0.04} \int_0^{2\pi} \int_0^{0.01} -5 \times 10^{-6} e^{-10^5 \rho z} \rho \, d\rho \, d\phi \, dz$$

We integrate first with respect to ϕ since it is so easy,

$$Q = \int_{0.02}^{0.04} \int_0^{0.01} -10^{-5} \pi e^{-10^5 \rho z} \rho \, d\rho \, dz$$

and then with respect to z, because this will simplify the last integration with respect to ρ,

$$Q = \int_0^{0.01} \left(\frac{-10^{-5}\pi}{-10^5 \rho} e^{-10^5 \rho z} \rho \, d\rho \right)_{z=0.02}^{z=0.04}$$

$$= \int_0^{0.01} -10^{-5} \pi (e^{-2000\rho} - e^{-4000\rho}) d\rho$$

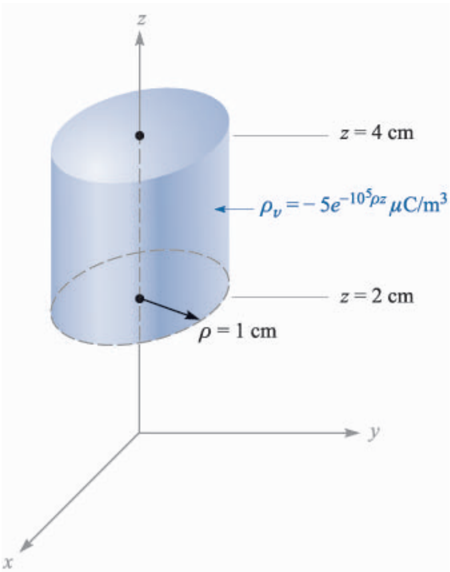


FIGURE 2.5
The total charge contained within the right circular cylinder may be obtained by evaluating

$$Q = \int_{\text{vol}} \rho_v \, dv.$$

Finally,

$$Q = -10^{-10}\pi\left(\frac{e^{-2000\rho}}{-2000} - \frac{e^{-4000\rho}}{-4000}\right)_0^{0.01}$$

$$Q = -10^{-10}\pi\left(\frac{1}{2000} - \frac{1}{4000}\right) = \frac{-\pi}{40} = 0.0785\,\text{pC}$$

where pC indicates picocoulombs.

Incidentally, we may use this result to make a rough estimate of the electron-beam current. If we assume these electrons are moving at a constant velocity of 10 percent of the velocity of light, this 2-cm-long packet will have moved 2 cm in $\frac{2}{3}$ ns, and the current is about equal to the product,

$$\frac{\Delta Q}{\Delta t} = \frac{-(\pi/40)10^{-12}}{(2/3)10^{-9}}$$

or approximately 118 μA.

The incremental contribution to the electric field intensity at $\mathbf{r}$ produced by an incremental charge ΔQ at $\mathbf{r}'$ is

$$\Delta\mathbf{E}(\mathbf{r}) = \frac{\Delta Q}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|^2}\,\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} = \frac{\rho_v\Delta v}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|^2}\,\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}$$

If we sum the contributions of all the volume charge in a given region and let the volume element Δv approach zero as the number of these elements becomes infinite, the summation becomes an integral,

$$\mathbf{E}(\mathbf{r}) = \int_{\text{vol}} \frac{\rho_v(\mathbf{r}')dv'}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|^2}\,\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} \tag{18}$$

This is again a triple integral, and (except in the drill problem that follows) we shall do our best to avoid actually performing the integration.

The significance of the various quantities under the integral sign of (18) might stand a little review. The vector $\mathbf{r}$ from the origin locates the field point where $\mathbf{E}$ is being determined, while the vector $\mathbf{r}'$ extends from the origin to the source point where $\rho_v(\mathbf{r}')dv'$ is located. The scalar distance between the source point and the field point is $|\mathbf{r}-\mathbf{r}'|$, and the fraction $(\mathbf{r}-\mathbf{r}')/|\mathbf{r}-\mathbf{r}'|$ is a unit vector directed from source point to field point. The variables of integration are x', y', and z' in cartesian coordinates.

✔ **D2.4.** Calculate the total charge within each of the indicated volumes: (*a*) $0.1 \le |x|, |y|, |z| \le 0.2$: $\rho_v = \frac{1}{x^3y^3z^3}$; (*b*) $0 \le \rho \le 0.1$, $0 \le \phi \le \pi$, $2 \le z \le 4$; $\rho_v = \rho^2z^2 \sin 0.6\phi$; (*c*) universe: $\rho_v = e^{-2r}/r^2$.

Ans. 27 μC; 1.778 mC; 6.28 C

2.4 FIELD OF A LINE CHARGE

Up to this point we have considered two types of charge distribution, the point charge and charge distributed throughout a volume with a density ρ_v C/m^3. If we now consider a filamentlike distribution of volume charge density, such as a very

fine, sharp beam in a cathode-ray tube or a charged conductor of very small radius, we find it convenient to treat the charge as a line charge of density ρ_L C/m. In the case of the electron beam the charges are in motion and it is true that we do not have an electrostatic problem. However, if the electron motion is steady and uniform (a dc beam) and if we ignore for the moment the magnetic field which is produced, the electron beam may be considered as composed of stationary electrons, for snapshots taken at any time will show the same charge distribution.

Let us assume a straight line charge extending along the z axis in a cylindrical coordinate system from $-\infty$ to ∞, as shown in Fig. 2.6. We desire the electric field intensity $\mathbf{E}$ at any and every point resulting from a *uniform* line charge density ρ_L.

Symmetry should always be considered first in order to determine two specific factors: (1) with which coordinates the field does *not* vary, and (2) which components of the field are *not* present. The answers to these questions then tell us which components are present and with which coordinates they *do* vary.

Referring to Fig. 2.6, we realize that as we move around the line charge, varying ϕ while keeping ρ and z constant, the line charge appears the same from every angle. In other words, azimuthal symmetry is present, and no field component may vary with ϕ.

Again, if we maintain ρ and ϕ constant while moving up and down the line charge by changing z, the line charge still recedes into infinite distance in both directions and the problem is unchanged. This is axial symmetry and leads to fields which are not functions of z.

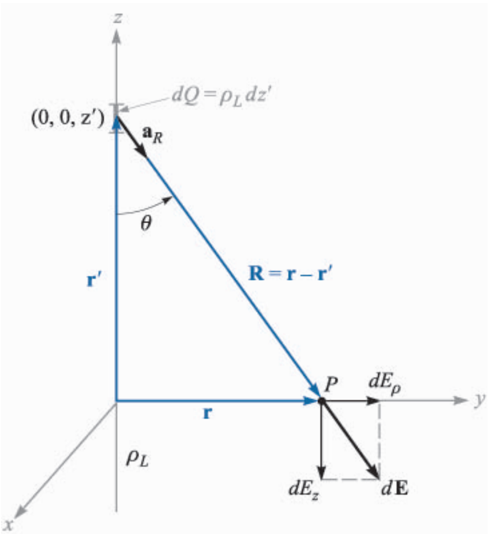


FIGURE 2.6
The contribution $d\mathbf{E} = dE_\rho \mathbf{a}_\rho + dE_z \mathbf{a}_z$ to the electric field intensity produced by an element of charge $dQ = \rho_L dz'$ located a distance z' from the origin. The linear charge density is uniform and extends along the entire z axis.

If we maintain ϕ and z constant and vary ρ, the problem changes, and Coulomb's law leads us to expect the field to become weaker as ρ increases. Hence, by a process of elimination we are led to the fact that the field varies only with ρ.

Now, which components are present? Each incremental length of line charge acts as a point charge and produces an incremental contribution to the electric field intensity which is directed away from the bit of charge (assuming a positive line charge). No element of charge produces a ϕ component of electric intensity; E_ϕ is zero. However, each element does produce an E_ρ and an E_z component, but the contribution to E_z by elements of charge which are equal distances above and below the point at which we are determining the field will cancel.

We therefore have found that we have only an E_ρ component and it varies only with ρ. Now to find this component.

We choose a point $P(0, y, 0)$ on the y axis at which to determine the field. This is a perfectly general point in view of the lack of variation of the field with ϕ and z. Applying (12) to find the incremental field at P due to the incremental charge $dQ = \rho_L dz'$, we have

$$d\mathbf{E} = \frac{\rho_L dz'(\mathbf{r} - \mathbf{r}')}{4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|^3}$$

where

$$\mathbf{r} = y\mathbf{a}_y = \rho\mathbf{a}_\rho$$
$$\mathbf{r}' = z'\mathbf{a}_z$$

and

$$\mathbf{r} - \mathbf{r}' = \rho\mathbf{a}_\rho - z'\mathbf{a}_z$$

Therefore,

$$d\mathbf{E} = \frac{\rho_L dz'(\rho\mathbf{a}_\rho - z'\mathbf{a}_z)}{4\pi\epsilon_0(\rho^2 + z'^2)^{3/2}}$$

Since only the $\mathbf{E}_\rho$ component is present, we may simplify:

$$dE_\rho = \frac{\rho_L \rho dz'}{4\pi\epsilon_0(\rho^2 + z'^2)^{3/2}}$$

and

$$E_\rho = \int_{-\infty}^{\infty} \frac{\rho_L \rho dz'}{4\pi\epsilon_0(\rho^2 + z'^2)^{3/2}}$$

Integrating by integral tables or change of variable, $z' = \rho\cot\theta$, we have

$$E_\rho = \frac{\rho_L}{4\pi\epsilon_0}\rho\left(\frac{1}{\rho^2}\frac{z'}{\sqrt{\rho^2 + z'^2}}\right)_{-\infty}^{\infty}$$

and

$$E_\rho = \frac{\rho_L}{2\pi\epsilon_0 \rho} \tag{19}$$

This is the desired answer, but there are many other ways of obtaining it. We might have used the angle θ as our variable of integration, for $z' = \rho \cot\theta$ from Fig. 2.6 and $dz' = -\rho \csc^2\theta \, d\theta$. Since $R = \rho \csc\theta$, our integral becomes, simply,

$$dE_\rho = \frac{\rho_L \, dz'}{4\pi\epsilon_0 R^2} \sin\theta = -\frac{\rho_L \sin\theta \, d\theta}{4\pi\epsilon_0 \rho}$$

$$E_\rho = -\frac{\rho_L}{4\pi\epsilon_0 \rho} \int_\pi^0 \sin\theta \, d\theta = \frac{\rho_L}{4\pi\epsilon_0 \rho} \cos\theta \Big]_\pi^0$$

$$= \frac{\rho_L}{2\pi\epsilon_0 \rho}$$

Here the integration was simpler, but some experience with problems of this type is necessary before we can unerringly choose the simplest variable of integration at the beginning of the problem.

We might also have considered (18) as our starting point,

$$\mathbf{E} = \int_{\text{vol}} \frac{\rho_v \, dv'(\mathbf{r} - \mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|^3}$$

letting $\rho_v \, dv' = \rho_L \, dz'$ and integrating along the line which is now our "volume" containing all the charge. Suppose we do this and forget everything we have learned from the symmetry of the problem. Choose point P now at a general location (ρ, ϕ, z) (Fig. 2.7) and write

$$\mathbf{r} = \rho \mathbf{a}_\rho + z \mathbf{a}_z$$

$$\mathbf{r}' = z' \mathbf{a}_z$$

$$\mathbf{R} = \mathbf{r} - \mathbf{r}' = \rho \mathbf{a}_\rho + (z - z') \mathbf{a}_z$$

$$R = \sqrt{\rho^2 + (z - z')^2}$$

$$\mathbf{a}_R = \frac{\rho \mathbf{a}_\rho + (z - z') \mathbf{a}_z}{\sqrt{\rho^2 + (z - z')^2}}$$

$$\mathbf{E} = \int_{-\infty}^{\infty} \frac{\rho_L \, dz' [\rho \mathbf{a}_\rho + (z - z') \mathbf{a}_z]}{4\pi\epsilon_0 [\rho^2 + (z - z')^2]^{3/2}}$$

$$= \frac{\rho_L}{4\pi\epsilon_0} \left\{ \int_{-\infty}^{\infty} \frac{\rho \, dz' \mathbf{a}_\rho}{[\rho^2 + (z - z')^2]^{3/2}} + \int_{-\infty}^{\infty} \frac{(z - z') \, dz' \mathbf{a}_z}{[\rho^2 + (z - z')^2]^{3/2}} \right\}$$

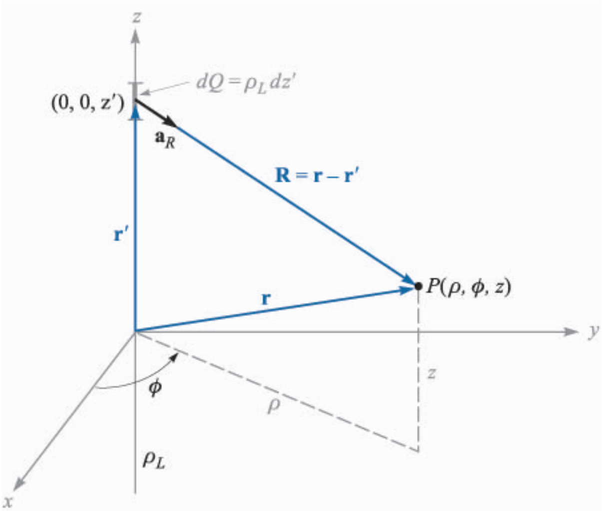


FIGURE 2.7
The geometry of the problem for the field about an infinite line charge leads to more difficult integrations when symmetry is ignored.

Before integrating a vector expression, we must know whether or not a vector under the integral sign (here the unit vectors $\mathbf{a}_\rho$ and $\mathbf{a}_z$) varies with the variable of integration (here dz'). If it does not, then it is a constant and may be removed from within the integral, leaving a scalar which may be integrated by normal methods. Our unit vectors, of course, cannot change in magnitude, but a change in direction is just as troublesome. Fortunately, the direction of $\mathbf{a}_\rho$ does not change with z' (nor with ρ, but it does change with ϕ), and $\mathbf{a}_z$ is constant always.

Hence we remove the unit vectors from the integrals and again integrate with tables or by changing variables,

$$\mathbf{E} = \frac{\rho_L}{4\pi\epsilon_0}\left\{\mathbf{a}_\rho \int_{-\infty}^{\infty} \frac{\rho\, dz'}{[\rho^2 + (z-z')^2]^{3/2}} + \mathbf{a}_z \int_{-\infty}^{\infty} \frac{(z-z')\, dz'}{[\rho^2 + (z-z')^2]^{3/2}}\right\}$$

$$= \frac{\rho_L}{4\pi\epsilon_0}\left\{\left[\mathbf{a}_\rho \rho \frac{1}{\rho^2} \frac{-(z-z')}{\sqrt{\rho^2 + (z-z')^2}}\right]_{-\infty}^{\infty} + \left[\mathbf{a}_z \frac{1}{\sqrt{\rho^2 + (z-z')^2}}\right]_{-\infty}^{\infty}\right\}$$

$$= \frac{\rho_L}{4\pi\epsilon_0}\left[\mathbf{a}_\rho \frac{2}{\rho} + \mathbf{a}_z(0)\right] = \frac{\rho_L}{2\pi\epsilon_0 \rho}\mathbf{a}_\rho$$

Again we obtain the same answer, as we should, for there is nothing wrong with the method except that the integration was harder and there were two integrations to perform. This is the price we pay for neglecting the consideration of symmetry and plunging doggedly ahead with mathematics. Look before you integrate.

Other methods for solving this basic problem will be discussed later after we introduce Gauss's law and the concept of potential.

Now let us consider the answer itself,

$$\mathbf{E} = \frac{\rho_L}{2\pi\epsilon_0 \rho}\mathbf{a}_\rho \tag{20}$$

We note that the field falls off inversely with the distance to the charged line, as compared with the point charge, where the field decreased with the *square* of the distance. Moving ten times as far from a point charge leads to a field only 1 percent the previous strength, but moving ten times as far from a line charge only reduces the field to 10 percent of its former value. An analogy can be drawn with a source of illumination, for the light intensity from a point source of light also falls off inversely as the square of the distance to the source. The field of an infinitely long fluorescent tube thus decays inversely as the first power of the radial distance to the tube, and we should expect the light intensity about a finite-length tube to obey this law near the tube. As our point recedes farther and farther from a finite-length tube, however, it eventually looks like a point source and the field obeys the inverse-square relationship.

Before leaving this introductory look at the field of the infinite line charge, we should recognize the fact that not all line charges are located along the z axis. As an example, let us consider an infinite line charge parallel to the z axis at $x = 6$, $y = 8$, Fig. 2.8. We wish to find $\mathbf{E}$ at the general field point $P(x, y, z)$.

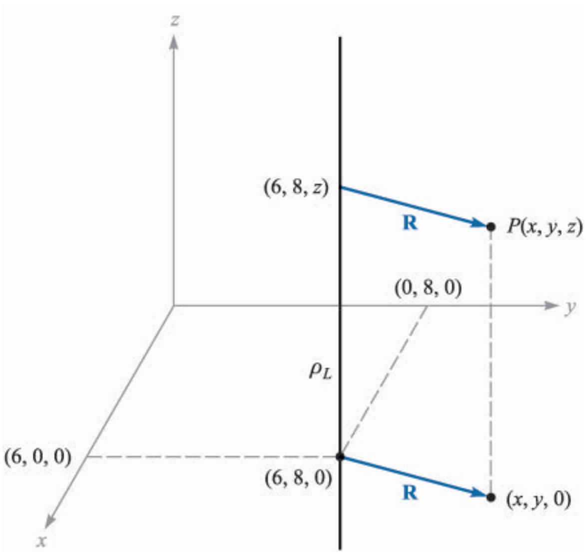


FIGURE 2.8
A point $P(x, y, z)$ is identified near an infinite uniform line charge located at $x = 6$, $y = 8$.

We replace ρ in (20) by the radial distance between the line charge and point, P, $R = \sqrt{(x-6)^2 + (y-8)^2}$, and let $\mathbf{a}_\rho$ be $\mathbf{a}_R$. Thus,

$$\mathbf{E} = \frac{\rho_L}{2\pi\epsilon_0\sqrt{(x-6)^2+(y-8)^2}}\mathbf{a}_R$$

where

$$\mathbf{a}_R = \frac{\mathbf{R}}{|\mathbf{R}|} = \frac{(x-6)\mathbf{a}_x + (y-8)\mathbf{a}_y}{\sqrt{(x-6)^2+(y-8)^2}}$$

Therefore,

$$\mathbf{E} = \frac{\rho_L}{2\pi\epsilon_0}\frac{(x-6)\mathbf{a}_x + (y-8)\mathbf{a}_y}{(x-6)^2+(y-8)^2}$$

We again note that the field is not a function of z.

In Sec. 2.6 we shall describe how fields may be sketched and use the field of the line charge as one example.

✔ **D2.5.** Infinite uniform line charges of 5 nC/m lie along the (positive and negative) x and y axes in free space. Find $\mathbf{E}$ at: (a) $P_A(0, 0, 4)$; (b) $P_B(0, 3, 4)$.

Ans. $45\mathbf{a}_z$ V/m; $10.8\mathbf{a}_y + 36.9\mathbf{a}_z$ V/m

2.5 FIELD OF A SHEET OF CHARGE

Another basic charge configuration is the infinite sheet of charge having a uniform density of ρ_S C/m^2. Such a charge distribution may often be used to approximate that found on the conductors of a strip transmission line or a parallel-plate capacitor. As we shall see in Chap. 5, static charge resides on conductor surfaces and not in their interiors; for this reason, ρ_S is commonly known as *surface charge density*. The charge-distribution family now is complete—point, line, surface, and volume, or Q, ρ_L, ρ_S, and ρ_v.

Let us place a sheet of charge in the yz plane and again consider symmetry (Fig. 2.9). We see first that the field cannot vary with y or with z, and then that the y and z components arising from differential elements of charge symmetrically located with respect to the point at which we wish the field will cancel. Hence only E_x is present, and this component is a function of x alone. We are again faced with a choice of many methods by which to evauate this component, and this time we shall use but one method and leave the others as exercises for a quiet Sunday afternoon.

Let us use the field of the infinite line charge (19) by dividing the infinite sheet into differential-width strips. One such strip is shown in Fig. 2.9. The line charge density, or charge per unit length, is $\rho_L = \rho_S\,dy'$, and the distance from

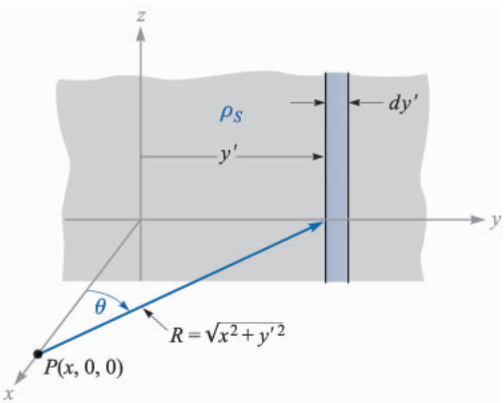


FIGURE 2.9
An infinite sheet of charge in the yz plane, a general point P on the x axis, and the differential-width line charge used as the element in determining the field at P by $d\mathbf{E} = \rho_S\, dy' \mathbf{a}_R/(2\pi\varepsilon_0 R)$.

this line charge to our general point P on the x axis is $R = \sqrt{x^2 + y'^2}$. The contribution to E_x at P from this differential-width strip is then

$$dE_x = \frac{\rho_S\, dy'}{2\pi\epsilon_0\sqrt{x^2 + y'^2}} \cos\theta = \frac{\rho_S}{2\pi\epsilon_0} \frac{x dy'}{x^2 + y'^2}$$

Adding the effects of all the strips,

$$E_x = \frac{\rho_S}{2\pi\epsilon_0}\int_{-\infty}^{\infty} \frac{x\, dy'}{x^2 + y'^2} = \frac{\rho_S}{2\pi\epsilon_0} \tan^{-1}\frac{y'}{x}\Big]_{-\infty}^{\infty} = \frac{\rho_S}{2\epsilon_0}$$

If the point P were chosen on the negative x axis, then

$$E_x = -\frac{\rho_S}{2\epsilon_0}$$

for the field is always directed away from the positive charge. This difficulty in sign is usually overcome by specifying a unit vector $\mathbf{a}_N$, which is normal to the sheet and directed outward, or away from it. Then

$$\boxed{\mathbf{E} = \frac{\rho_S}{2\epsilon_0}\mathbf{a}_N} \tag{21}$$

This is a startling answer, for the field is constant in magnitude and direction. It is just as strong a million miles away from the sheet as it is right off the surface. Returning to our light analogy, we see that a uniform source of light on the ceiling of a very large room leads to just as much illumination on a square foot on the floor as it does on a square foot a few inches below the ceiling. If you desire greater illumination on this subject, it will do you no good to hold the book closer to such a light source.

If a second infinite sheet of charge, having a *negative* charge density $-\rho_S$, is located in the plane $x = a$, we may find the total field by adding the contribution of each sheet. In the region $x > a$,

$$\mathbf{E}_+ = \frac{\rho_S}{2\epsilon_0}\mathbf{a}_x \qquad \mathbf{E}_- = -\frac{\rho_S}{2\epsilon_0}\mathbf{a}_x \qquad \mathbf{E} = \mathbf{E}_+ + \mathbf{E}_- = 0$$

and for $x < 0$,

$$\mathbf{E}_+ = -\frac{\rho_S}{2\epsilon_0}\mathbf{a}_x \qquad \mathbf{E}_- = \frac{\rho_S}{2\epsilon_0}\mathbf{a}_x \qquad \mathbf{E} = \mathbf{E}_+ + \mathbf{E}_- = 0$$

and when $0 < x < a$,

$$\mathbf{E}_+ = \frac{\rho_S}{2\epsilon_0}\mathbf{a}_x \qquad \mathbf{E}_- = \frac{\rho_S}{2\epsilon_0}\mathbf{a}_x$$

and

$$\boxed{\mathbf{E} = \mathbf{E}_+ + \mathbf{E}_- = \frac{\rho_S}{\epsilon_0}\mathbf{a}_x} \qquad (22)$$

This is an important practical answer, for it is the field between the parallel plates of an air capacitor, provided the linear dimensions of the plates are very much greater than their separation and provided also that we are considering a point well removed from the edges. The field outside the capacitor, while not zero, as we found for the ideal case above, is usually negligible.

✔ **D2.6.** Three infinite uniform sheets of charge are located in free space as follows: 3 nC/m^2 at $z = -4$, 6 nC/m^2 at $z = 1$, and -8 nC/m^2 at $z = 4$. Find **E** at the point: (*a*) $P_A(2, 5, -5)$; (*b*) $P_B(4, 2, -3)$; (*c*) $P_C(-1, -5, 2)$; (*d*) $P_D(-2, 4, 5)$.

Ans. $-56.5\mathbf{a}_z$; $283\mathbf{a}_z$; $961\mathbf{a}_z$; $56.5\mathbf{a}_z$ all V/m

2.6 STREAMLINES AND SKETCHES OF FIELDS

We now have vector equations for the electric field intensity resulting from several different charge configurations, and we have had little difficulty in interpreting the magnitude and direction of the field from the equations. Unfortunately, this simplicity cannot last much longer, for we have solved most of the simple cases and our new charge distributions must lead to more complicated expressions for the fields and more difficulty in visualizing the fields through the equations. However, it is true that one picture would be worth about a thousand words, if we just knew what picture to draw.

Consider the field about the line charge,

$$\mathbf{E} = \frac{\rho_L}{2\pi\epsilon_0\rho}\mathbf{a}_\rho$$

Fig. 2.10*a* shows a cross-sectional view of the line charge and presents what might be our first effort at picturing the field—short line segments drawn here and there having lengths proportional to the magnitude of **E** and pointing in the direction of **E**. The figure fails to show the symmetry with respect to ϕ, so we try again in Fig. 2.10*b* with a symmetrical location of the line segments. The real trouble now appears—the longest lines must be drawn in the most crowded region, and this also plagues us if we use line segments of equal length but of a thickness which is proportional to **E** (Fig. 2.10*c*). Other schemes which have been suggested include drawing shorter lines to represent stronger fields (inherently misleading) and using intensity of color to represent stronger fields (difficult and expensive).

For the present, then, let us be content to show only the *direction* of **E** by drawing continuous lines from the charge which are everywhere tangent to **E**. Fig. 2.10*d* shows this compromise. A symmetrical distribution of lines (one every 45°) indicates azimuthal symmetry, and arrowheads should be used to show direction.

These lines are usually called *streamlines*, although other terms such as flux lines and direction lines are also used. A small positive test charge placed at any

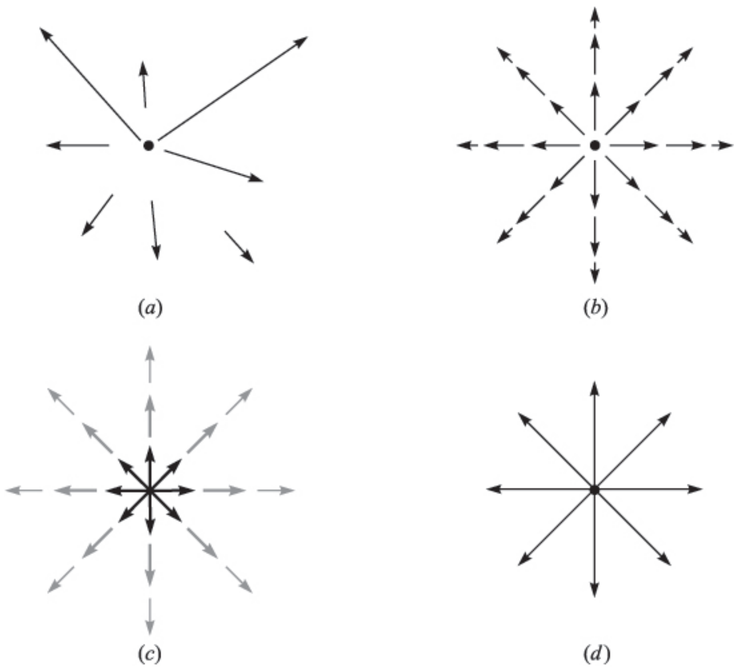


FIGURE 2.10
(*a*) One very poor sketch, (*b*) and (*c*) two fair sketches, and (*d*) the usual form of streamline sketch. In the last form, the arrows show the direction of the field at every point along the line, and the spacing of the lines is inversely proportional to the strength of the field.

point in this field and free to move would accelerate in the direction of the streamline passing through that point. If the field represented the velocity of a liquid or a gas (which, incidentally, would have to have a source at $\rho = 0$), small suspended particles in the liquid or gas would trace out the streamlines.

We shall find out later that a bonus accompanies this streamline sketch, for the magnitude of the field can be shown to be inversely proportional to the spacing of the streamlines for some important special cases. The closer they are together, the stronger is the field. At that time we shall also find an easier, more accurate method of making that type of streamline sketch.

If we attempted to sketch the field of the point charge, the variation of the field into and away from the page would cause essentially insurmountable difficulties; for this reason sketching is usually limited to two-dimensional fields.

In the case of the two-dimensional field let us arbitrarily set $E_z = 0$. The streamlines are thus confined to planes for which z is constant, and the sketch is the same for any such plane. Several streamlines are shown in Fig. 2.11, and the E_x and E_y components are indicated at a general point. Since it is apparent from the geometry that

$$\frac{E_y}{E_x} = \frac{dy}{dx} \tag{23}$$

a knowledge of the functional form of E_x and E_y (and the ability to solve the resultant differential equation) will enable us to obtain the equations of the streamlines.

As an illustration of this method, consider the field of the uniform line charge with $\rho_L = 2\pi\epsilon_0$,

$$\mathbf{E} = \frac{1}{\rho}\mathbf{a}_\rho$$

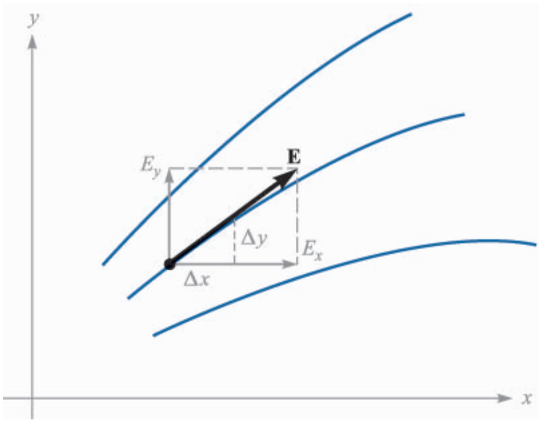


FIGURE 2.11
The equation of a streamline is obtained by solving the differential equation $E_y/E_x = dy/dx$.

In cartesian coordinates,

$$\mathbf{E} = \frac{x}{x^2 + y^2}\mathbf{a}_x + \frac{y}{x^2 + y^2}\mathbf{a}_y$$

Thus we form the differential equation

$$\frac{dy}{dx} = \frac{E_y}{E_x} = \frac{y}{x} \qquad \text{or} \qquad \frac{dy}{y} = \frac{dx}{x}$$

Therefore,

$$\ln y = \ln x + C_1 \qquad \text{or} \qquad \ln y = \ln x + \ln C$$

from which the equations of the streamlines are obtained,

$$y = Cx$$

If we want to find the equation of one particular streamline, say that one passing through $P(-2, 7, 10)$, we merely substitute the coordinates of that point into our equation and evaluate C. Here, $7 = C(-2)$, and $C = -3.5$, so that $y = -3.5x$.

Each streamline is associated with a specific value of C, and the radial lines shown in Fig. 2.10*d* are obtained when $C = 0, 1, -1$, and $1/C = 0$.

The equations of streamlines may also be obtained directly in cylindrical or spherical coordinates. A spherical coordinate example will be examined in Sec. 4.7.

✔ **D2.7.** Find the equation of that streamline that passes through the point $P(1, 4, -2)$ in the field $\mathbf{E} =$: (*a*) $\frac{-8x}{y}\mathbf{a}_x + \frac{4x^2}{y^2}\mathbf{a}_y$; (*b*) $2e^{5x}[y(5x+1)\mathbf{a}_x + x\mathbf{a}_y]$.

Ans. $x^2 + 2y^2 = 33$; $y^2 = 15.6 + 0.4x - 0.08\ln(x + 0.2)/1.2]$

SUGGESTED REFERENCES

1. Boast, W. B.: "Vector Fields," Harper and Row, Publishers, Incorporated, New York, 1964. This book contains numerous examples and sketches of fields.
2. Della Torre, E., and Longo, C. L.: "The Electromagnetic Field," Allyn and Bacon, Inc., Boston, 1969. The authors introduce all of electromagnetic theory with a careful and rigorous development based on a single experimental law—that of Coulomb. It begins in chap. 1.
3. Schelkunoff, S. A.: "Electromagnetic Fields," Blaisdell Publishing Company, New York, 1963. Many of the physical aspects of fields are discussed early in this text without advanced mathematics.

PROBLEMS

2.1 Four 10-nC positive charges are located in the $z = 0$ plane at the corners of a square 8 cm on a side. A fifth 10-nC positive charge is located at a point 8 cm distant from each of the other charges. Calculate the magnitude of the total force on this fifth charge for $\epsilon = \epsilon_0$.

2.2 A charge $Q_1 = 0.1\,\mu C$ is located at the origin in free space, while $Q_2 = 0.2\,\mu C$ is at $A(0.8, -0.6, 0)$. Find the locus of points in the $z = 0$ plane at which the x-component of the force on a third positive charge is zero.

2.3 Point charges of 50 nC each are located at $A(1, 0, 0)$, $B(-1, 0, 0)$, $C(0, 1, 0)$, and $D(0, -1, 0)$ in free space. Find the total force on the charge at A.

2.4 Let $Q_1 = 8\,\mu C$ be located at $P_1(2, 5, 8)$ while $Q_2 = -5\,\mu C$ is at $P_2(6, 15, 8)$. Let $\epsilon = \epsilon_0$. (*a*) Find $\mathbf{F}_2$, the force on Q_2. (*b*) Find the coordinates of P_3 if a charge Q_3 experiences a total force $\mathbf{F}_3 = 0$ at P_3.

2.5 Let a point charge $Q_1 = 25$ nC be located at $P_1(4, -2, 7)$ and a charge $Q_2 = 60$ nC be at $P_2(-3, 4, -2)$. (*a*) If $\epsilon = \epsilon_0$, find $\mathbf{E}$ at $P(1, 2, 3)$. (*b*) At what point on the y axis is $E_x = 0$?

2.6 Point charges of 120 nC are located at $A(0, 0, 1)$ and $B(0, 0, -1)$ in free space. (*a*) Find $\mathbf{E}$ at $P(0.5, 0, 0)$. (*b*) What single charge at the origin would provide the identical field strength?

2.7 A 2-μC point charge is located at $A(4, 3, 5)$ in free space. Find E_ρ, E_ϕ, and E_z at $P(8, 12, 2)$.

2.8 Given point charges of $-1\,\mu C$ at $P_1(0, 0, 0.5)$ and $P_2(0, 0, -0.5)$, and a charge of $2\,\mu C$ at the origin, find $\mathbf{E}$ at $P(0, 2, 1)$ in spherical components. Assume $\epsilon = \epsilon_0$.

2.9 A 100-nC point charge is located at $A(-1, 1, 3)$ in free space. (*a*) Find the locus of all points $P(x, y, z)$ at which $E_x = 500$ V/m. (*b*) Find y_1 if $P(-2, y_1, 3)$ lies on that locus.

2.10 Charges of 20 and -20 nC are located at $(3, 0, 0)$ and $(-3, 0, 0)$, respectively. Let $\epsilon = \epsilon_0$. (*a*) Determine $|\mathbf{E}|$ at $P(0, y, 0)$. (*b*) Sketch $|\mathbf{E}|$ vs y at P.

2.11 A charge Q_0, located at the origin in free space, produces a field for which $E_z = 1$ kV/m at point $P(-2, 1, -1)$. (*a*) Find Q_0. Find $\mathbf{E}$ at $M(1, 6, 5)$ in: (*b*) cartesian coordinates; (*c*) cylindrical coordinates; (*d*) spherical coordinates.

2.12 The volume charge density $\rho_v = \rho_0 e^{-|x|-|y|-|z|}$ exists over all free space. Calculate the total charge present.

2.13 A uniform volume charge density of $0.2\,\mu C/m^2$ is present throughout the spherical shell extending from $r = 3$ cm to $r = 5$ cm. If $\rho_v = 0$ elsewhere, find: (*a*) the total charge present within the shell, and (*b*) r_1 if half the total charge is located in the region 3 cm $< r < r_1$.

2.14 Let $\rho_v = 5e^{-0.1\rho}(\pi - |\phi|)\dfrac{1}{z^2 + 10}\,\mu C/m^3$ in the region $0 \le \rho \le 10$, $-\pi < \phi < \pi$, all z, and $\rho_v = 0$ elsewhere. (*a*) Determine the total charge

present. (*b*) Calculate the charge within the region $0 \le \rho \le 4$, $-\frac{1}{2}\pi < \phi < \frac{1}{2}\pi$, $-10 < z < 10$.

2.15 A spherical volume having a 2-μm radius contains a uniform volume charge density of $10^{15}\ \mathrm{C/m^3}$. (*a*) What total charge is enclosed in the spherical volume? (*b*) Now assume that a large region contains one of these little spheres at every corner of a cubical grid 3 mm on a side, and that there is no charge between the spheres. What is the average volume charge density throughout this large region?

2.16 The region in which $4 < r < 5$, $0 < \theta < 25^\circ$, and $0.9\pi < \phi < 1.1\pi$, contains the volume charge density $\rho_v = 10(r-4)(r-5)\sin\theta\sin\frac{1}{2}\phi$. Outside that region $\rho_v = 0$. Find the charge within the region.

2.17 A uniform line charge of 16 nC/m is located along the line defined by $y = -2$, $z = 5$. If $\epsilon = \epsilon_0$: (*a*) find **E** at $P(1, 2, 3)$; (*b*) find **E** at that point in the $z = 0$ plane where the direction of **E** is given by $\frac{1}{3}\mathbf{a}_y - \frac{2}{3}\mathbf{a}_z$.

2.18 Uniform line charges of 0.4 μC/m and −0.4 μC/m are located in the $x = 0$ plane at $y = -0.6$ and $y = 0.6$ m, respectively. Let $\epsilon = \epsilon_0$. Find **E** at: (*a*) $P(x, 0, z)$; (*b*) $Q(2, 3, 4)$.

2.19 A uniform line charge of 2 μC/m is located on the z axis. Find **E** in cartesian coordinates at $P(1, 2, 3)$ if the charge extends from: (*a*) $z = -\infty$ to $z = \infty$; (*b*) $z = -4$ to $z = 4$.

2.20 Uniform line charges of 120 nC/m lie along the entire extent of the three coordinate axes. Assuming free space conditions, find **E** at $P(-3, 2, -1)$.

2.21 Two identical uniform line charges, with $\rho_L = 75$ nC/m, are located in free space at $x = 0$, $y = \pm 0.4$ m. What force per unit length does each line charge exert on the other?

2.22 A uniform surface charge density of $5\ \mathrm{nC/m^2}$ is present in the region $x = 0$, $-2 < y < 2$, all z. If $\epsilon = \epsilon_0$, find **E** at: (*a*) $P_A(3, 0, 0)$; (*b*) $P_B(0, 3, 0)$.

2.23 Given the surface charge density, $\rho_S = 2\ \mu\mathrm{C/m^2}$ in the region $\rho < 0.2$ m, $z = 0$, and is zero elsewhere, find **E** at: (*a*) $P_A(\rho = 0, z = 0.5)$; (*b*) $P_B(\rho = 0, z = -0.5)$.

2.24 Surface charge density is positioned in free space as follows: $20\ \mathrm{nC/m^2}$ at $x = -3$, $-30\ \mathrm{nC/m^2}$ at $y = 4$, and $40\ \mathrm{nC/m^2}$ at $z = 2$. Find the magnitude of **E** at: (*a*) $P_A(4, 3, -2)$; (*b*) $P_B(-2, 5, -1)$; (*c*) $P_C(0, 0, 0)$.

2.25 Find **E** at the origin if the following charge distributions are present in free space: point charge, 12 nC, at $P(2, 0, 6)$; uniform line charge density, 3 nC/m, at $x = -2$, $y = 3$; uniform surface charge density, $0.2\ \mathrm{nC/m^2}$, at $x = 2$.

2.26 A uniform line charge density of 5 nC/m is at $y = 0$, $z = 2$ m in free space, while −5 nC/m is located at $y = 0$, $z = -2$ m. A uniform surface charge density of $0.3\ \mathrm{nC/m^2}$ is at $y = 0.2$ m, and $-0.3\ \mathrm{nC/m^2}$ is at $y = -0.2$ m. Find $|\mathbf{E}|$ at the origin.

2.27 Given the electric field $\mathbf{E} = (4x - 2y)\mathbf{a}_x - (2x + 4y)\mathbf{a}_y$, find: (*a*) the equation of that streamline passing through the point $P(2, 3, -4)$; (*b*) a unit vector $\mathbf{a}_E$ specifying the direction of **E** at $Q(3, -2, 5)$.

2.28 Let $\mathbf{E} = 5x^3\mathbf{a}_x - 15x^2y\mathbf{a}_y$, and find: (*a*) the equation of the streamline that passes through $P(4, 2, 1)$; (*b*) a unit vector $\mathbf{a}_E$ specifying the direction of $\mathbf{E}$ at $Q(3, -2, 5)$; (*c*) a unit vector $\mathbf{a}_N = (l, m, 0)$ that is perpendicular to $\mathbf{a}_E$ at Q.

2.29 If $\mathbf{E} = 20e^{-5y}(\cos 5x\mathbf{a}_x - \sin 5x\mathbf{a}_y)$, find: (*a*) $|\mathbf{E}|$ at $P(\pi/6, 0.1, 2)$; (*b*) a unit vector in the direction of $\mathbf{E}$ at P; (*c*) the equation of the direction line passing through P.

2.30 Given the electric field intensity, $\mathbf{E} = 400y\mathbf{a}_x + 400x\mathbf{a}_y$ V/m, find: (*a*) the equation of the streamline passing through point $A(2, 1, -2)$; (*b*) the equation of the surface on which $|\mathbf{E}| = 800$ V. (*c*) Sketch the streamline of part *a*. (*d*) Sketch the trace produced by the intersection of the $z = 0$ plane and the surface of part *b*.

2.31 In cylindrical coordinates with $\mathbf{E}(\rho, \phi) = \mathbf{E}_\rho(\rho, \phi)\mathbf{a}_\rho + E_\phi(\rho, \phi)\mathbf{a}_\phi$, the differential equation describing the direction lines is $E_\rho/E_\phi = d\rho/(\rho\, d\phi)$ in any $z =$ constant plane. Derive the equation of the line passing through point $P(\rho = 4, \phi = 10^\circ, z = 2)$ in the field $\mathbf{E} = 2\rho^2 \cos 3\phi\mathbf{a}_\rho + 2\rho^2 \sin 3\phi\mathbf{a}_\phi$.

CHAPTER 3

ELECTRIC FLUX DENSITY, GAUSS'S LAW, AND DIVERGENCE

After drawing a few of the fields described in the previous chapter and becoming familiar with the concept of the streamlines which show the direction of the force on a test charge at every point, it is difficult to avoid giving these lines a physical significance and thinking of them as *flux* lines. No physical particle is projected radially outward from the point charge, and there are no steel tentacles reaching out to attract or repel an unwary test charge, but as soon as the streamlines are drawn on paper there seems to be a picture showing "something" is present.

It is very helpful to invent an *electric flux* which streams away symmetrically from a point charge and is coincident with the streamlines and to visualize this flux wherever an electric field is present.

This chapter introduces and uses the concept of electric flux and electric flux density to solve again several of the problems presented in the last chapter. The work here turns out to be much easier, and this is due to the extremely symmetrical problems which we are solving.

3.1 ELECTRIC FLUX DENSITY

About 1837 the Director of the Royal Society in London, Michael Faraday, became very interested in static electric fields and the effect of various insulating materials on these fields. This problem had been bothering him during the past ten years when he was experimenting in his now famous work on induced electromotive force, which we shall discuss in Chap. 10. With that subject completed, he had a pair of concentric metallic spheres constructed, the outer one consisting of two hemispheres that could be firmly clamped together. He also prepared shells of insulating material (or dielectric material, or simply *dielectric*) which would occupy the entire volume between the concentric spheres. We shall not make immediate use of his findings about dielectric materials, for we are restricting our attention to fields in free space until Chap. 5. At that time we shall see that the materials he used will be classified as ideal dielectrics.

His experiment, then, consisted essentially of the following steps:

1. With the equipment dismantled, the inner sphere was given a known positive charge.
2. The hemispheres were then clamped together around the charged sphere with about 2 cm of dielectric material between them.
3. The outer sphere was discharged by connecting it momentarily to ground.
4. The outer space was separated carefully, using tools made of insulating material in order not to disturb the induced charge on it, and the negative induced charge on each hemisphere was measured.

Faraday found that the total charge on the outer sphere was equal in *magnitude* to the original charge placed on the inner sphere and that this was true regardless of the dielectric material separating the two spheres. He concluded that there was some sort of "displacement" from the inner sphere to the outer which was independent of the medium, and we now refer to this flux as *displacement*, *displacement flux*, or simply *electric flux*.

Faraday's experiments also showed, of course, that a larger positive charge on the inner sphere induced a correspondingly larger negative charge on the outer sphere, leading to a direct proportionality between the electric flux and the charge on the inner sphere. The constant of proportionality is dependent on the system of units involved, and we are fortunate in our use of SI units, because the constant is unity. If electric flux is denoted by Ψ (psi) and the total charge on the inner sphere by Q, then for Faraday's experiment

$$\boxed{\Psi = Q}$$

and the electric flux Ψ is measured in coulombs.

We can obtain more quantitative information by considering an inner sphere of radius a and an outer sphere of radius b, with charges of Q and $-Q$,

respectively (Fig. 3.1). The paths of electric flux Ψ extending from the inner sphere to the outer sphere are indicated by the symmetrically distributed streamlines drawn radially from one sphere to the other.

At the surface of the inner sphere, Ψ coulombs of electric flux are produced by the charge $Q(=\Psi)$ coulombs distributed uniformly over a surface having an area of $4\pi a^2\,\text{m}^2$. The density of the flux at this surface is $\Psi/4\pi a^2$ or $Q/4\pi a^2\,\text{C/m}^2$, and this is an important new quantity.

Electric flux density, measured in coulombs per square meter (sometimes described as "lines per square meter," for each line is due to one coulomb), is given the letter **D**, which was originally chosen because of the alternate names of *displacement flux density* or *displacement density*. Electric flux density is more descriptive, however, and we shall use the term consistently.

The electric flux density **D** is a vector field and is a member of the "flux density" class of vector fields, as opposed to the "force fields" class, which includes the electric field intensity **E**. The direction of **D** at a point is the direction of the flux lines at that point, and the magnitude is given by the number of flux lines crossing a surface normal to the lines divided by the surface area.

Referring again to Fig. 3.1, the electric flux density is in the radial direction and has a value of

$$\mathbf{D}\Big|_{r=a} = \frac{Q}{4\pi a^2}\mathbf{a}_r \qquad \text{(inner sphere)}$$

$$\mathbf{D}\Big|_{r=b} = \frac{Q}{4\pi b^2}\mathbf{a}_r \qquad \text{(outer sphere)}$$

and at a radial distance r, where $a \le r \le b$,

$$\mathbf{D} = \frac{Q}{4\pi r^2}\mathbf{a}_r$$

If we now let the inner sphere become smaller and smaller, while still retaining a charge of Q, it becomes a point charge in the limit, but the electric flux density at a point r meters from the point charge is still given by

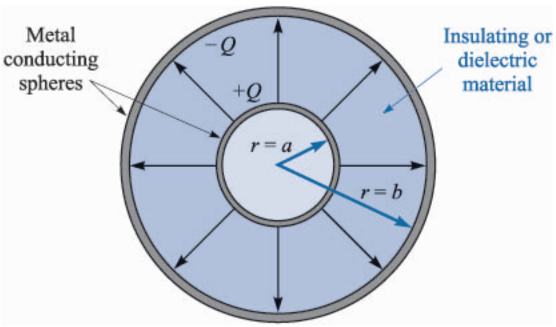


FIGURE 3.1
The electric flux in the region between a pair of charged concentric spheres. The direction and magnitude of **D** are not functions of the dielectric between the spheres.

$$\mathbf{D} = \frac{Q}{4\pi r^2}\mathbf{a}_r \tag{1}$$

for Q lines of flux are symmetrically directed outward from the point and pass through an imaginary spherical surface of area $4\pi r^2$.

This result should be compared with Sec. 2.2, Eq. (10), the radial electric field intensity of a point charge in free space,

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2}\mathbf{a}_r$$

In free space, therefore,

$$\mathbf{D} = \epsilon_0\mathbf{E} \qquad \text{(free space only)} \tag{2}$$

Although (2) is applicable only to a vacuum, it is not restricted solely to the field of a point charge. For a general volume charge distribution in free space

$$\mathbf{E} = \int_{\text{vol}} \frac{\rho_v dv}{4\pi\epsilon_0 R^2}\mathbf{a}_r \qquad \text{(free space only)} \tag{3}$$

where this relationship was developed from the field of a single point charge. In a similar manner, (1) leads to

$$\mathbf{D} = \int_{\text{vol}} \frac{\rho_v\, dv}{4\pi R^2}\mathbf{a}_r \tag{4}$$

and (2) is therefore true for any free-space charge configuration; we shall consider (2) as defining **D** in free space.

As a preparation for the study of dielectrics later, it might be well to point out now that, for a point charge embedded in an infinite ideal dielectric medium, Faraday's results show that (1) is still applicable, and thus so is (4). Equation (3) is not applicable, however, and so the relationship between **D** and **E** will be slightly more complicated than (2).

Since **D** is directly proportional to **E** in free space, it does not seem that it should really be necessary to introduce a new symbol. We do so for several reasons. First, **D** is associated with the flux concept, which is an important new idea. Second, the **D** fields we obtain will be a little simpler than the corresponding **E** fields, since ϵ_0 does not appear. And, finally, it helps to become a little familiar with **D** before it is applied to dielectric materials in Chap. 5.

Let us consider a simple numerical example to illustrate these new quantities and units.

Example 3.1

We wish to find **D** in the region about a uniform line charge of 8 nC/m lying along the z axis in free space.

Solution. The **E** field is

$$\mathbf{E} = \frac{\rho_L}{2\pi\epsilon_0\rho}\mathbf{a}_\rho = \frac{8 \times 10^{-9}}{2\pi(8.854 \times 10^{-12})\rho}\mathbf{a}_\rho = \frac{143.8}{\rho}\mathbf{a}_\rho \text{ V/m}$$

At $\rho = 3$ m, $\mathbf{E} = 47.9\mathbf{a}_\rho$ V/m.

Associated with the **E** field, we find

$$\mathbf{D} = \frac{\rho_L}{2\pi\rho}\mathbf{a}_\rho = \frac{8 \times 10^{-9}}{2\pi\rho}\mathbf{a}_\rho = \frac{1.273 \times 10^{-9}}{\rho}\mathbf{a}_\rho \text{ C/m}^2$$

The value at $\rho = 3$ m is $\mathbf{D} = 0.424\mathbf{a}_\rho$ nC/m.

The total flux leaving a 5-m length of the line charge is equal to the total charge on that length, or $\Psi = 40$ nC.

✔ **D3.1.** Given a 60-μC point charge located at the origin, find the total electric flux passing through: (a) that portion of the sphere $r = 26$ cm bounded by $0 < \theta < \frac{\pi}{2}$ and $0 < \phi < \frac{\pi}{2}$; (b) the closed surface defined by $\rho = 26$ cm and $z = \pm 26$ cm; (c) the plane $z = 26$ cm.

Ans. 7.5 μC; 60 μC; 30 μC

✔ **D3.2.** Calculate **D** in rectangular coordinates at point $P(2, -3, 6)$ produced by: (a) a point charge $Q_A = 55$ mC at $Q(-2, 3, -6)$; (b) a uniform line charge $\rho_{LB} = 20$ mC/m on the x axis; (c) a uniform surface charge density $\rho_{SC} = 120\ \mu\text{C/m}^2$ on the plane $z = -5$ m.

Ans. $6.38\mathbf{a}_x - 9.57\mathbf{a}_y + 19.14\mathbf{a}_z\ \mu\text{C/m}^2$; $-212\mathbf{a}_y + 424\mathbf{a}_z\ \mu\text{C/m}^2$; $60\mathbf{a}_z\ \mu\text{C/m}^2$

3.2 GAUSS'S LAW

The results of Faraday's experiments with the concentric spheres could be summed up as an experimental law by stating that the electric flux passing through any imaginary spherical surface lying between the two conducting spheres is equal to the charge enclosed within that imaginary surface. This enclosed charge is distributed on the surface of the inner sphere, or it might be concentrated as a point charge at the center of the imaginary sphere. However, since one coulomb of electric flux is produced by one coulomb of charge, the inner conductor might just as well have been a cube or a brass door key and the total induced charge on the outer sphere would still be the same. Certainly the flux density would change from its previous symmetrical distribution to some

unknown configuration, but $+Q$ coulombs on any inner conductor would produce an induced charge of $-Q$ coulombs on the surrounding sphere. Going one step further, we could now replace the two outer hemispheres by an empty (but completely closed) soup can. Q coulombs on the brass door key would produce $\Psi = Q$ lines of electric flux and would induce $-Q$ coulombs on the tin can.[1]

These generalizations of Faraday's experiment lead to the following statement, which is known as *Gauss's law*:

> The electric flux passing through any closed surface is equal to the total charge enclosed by that surface.

The contribution of Gauss, one of the greatest mathematicians the world has ever produced, was actually not in stating the law as we have above, but in providing a mathematical form for this statement, which we shall now obtain.

Let us imagine a distribution of charge, shown as a cloud of point charges in Fig. 3.2, surrounded by a closed surface of any shape. The closed surface may be the surface of some real material, but more generally it is any closed surface we wish to visualize. If the total charge is Q, then Q coulombs of electric flux will pass through the enclosing surface. At every point on the surface the electric-flux-density vector $\mathbf{D}$ will have some value $\mathbf{D}_S$, where the subscript S merely reminds us that $\mathbf{D}$ must be evaluated at the surface, and $\mathbf{D}_S$ will in general vary in magnitude and direction from one point on the surface to another.

We must now consider the nature of an incremental element of the surface. An incremental element of area ΔS is very nearly a portion of a plane surface, and the complete description of this surface element requires not only a statement of its magnitude ΔS but also of its orientation in space. In other words, the incremental surface element is a vector quantity. The only unique direction which may be associated with $\Delta \mathbf{S}$ is the direction of the normal to that plane which is tangent to the surface at the point in question. There are, of course, two

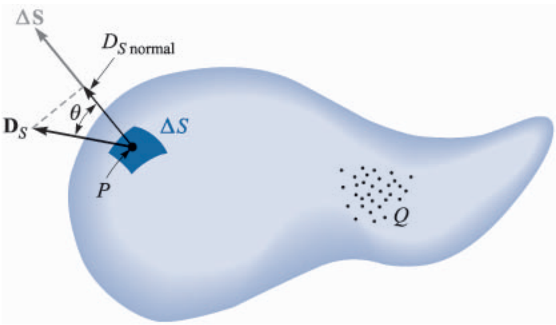


FIGURE 3.2
The electric flux density $\mathbf{D}_S$ at P due to charge Q. The total flux passing through $\Delta\mathbf{S}$ is $\mathbf{D}_S \cdot \Delta\mathbf{S}$.

[1] If it were a perfect insulator, the soup could even be left in the can without any difference in the results.

such normals, and the ambiguity is removed by specifying the outward normal whenever the surface is closed and "outward" has a specific meaning.

At any point P consider an incremental element of surface ΔS and let $\mathbf{D}_S$ make an angle θ with ΔS, as shown in Fig. 3.2. The flux crossing $\Delta\mathbf{S}$ is then the product of the normal component of $\mathbf{D}_S$ and $\Delta\mathbf{S}$,

$$\Delta\Psi = \text{flux crossing } \Delta\mathbf{S} = D_{S,norm}\Delta = D_S S\cos\theta\Delta S = \mathbf{D}_S \cdot \Delta\mathbf{S}$$

where we are able to apply the definition of the dot product developed in Chap. 1.

The *total* flux passing through the closed surface is obtained by adding the differential contributions crossing each surface element $\Delta\mathbf{S}$,

$$\Psi = \int d\Psi = \oint_{\substack{\text{closed}\\ \text{surface}}} \mathbf{D}_S \cdot d\mathbf{S}$$

The resultant integral is a *closed surface integral*, and since the surface element $d\mathbf{S}$ always involves the differentials of two coordinates, such as $dx\,dy$, $\rho\,d\phi\,d\rho$, or $r^2\sin\theta\,d\theta\,d\phi$, the integral is a double integral. Usually only one integral sign is used for brevity, and we shall always place an S below the integral sign to indicate a surface integral, although this is not actually necessary since the differential $d\mathbf{S}$ is automatically the signal for a surface integral. One last convention is to place a small circle on the integral sign itself to indicate that the integration is to be performed over a *closed* surface. Such a surface is often called a *gaussian surface*. We then have the mathematical formulation of Gauss's law,

$$\boxed{\Psi = \oint_S \mathbf{D}_S \cdot d\mathbf{S} = \text{charge enclosed} = Q} \qquad (5)$$

The charge enclosed might be several point charges, in which case

$$Q = \Sigma Qn$$

or a line charge,

$$Q = \int \rho_L\,dL$$

or a surface charge,

$$Q = \int_S \rho_S dS \qquad \text{(not necessarily a closed surface)}$$

or a volume charge distribution,

$$Q = \int_{\text{vol}} \rho_v\,dv$$

The last form is usually used, and we should agree now that it represents any or all of the other forms. With this understanding Gauss's law may be written in terms of the charge distribution as

$$\oint_S \mathbf{D}_S \cdot d\mathbf{S} = \int_{\text{vol}} \rho_v \, dv \tag{6}$$

a mathematical statement meaning simply that the total electric flux through any closed surface is equal to the charge enclosed.

To illustrate the application of Gauss's law, let us check the results of Faraday's experiment by placing a point charge Q at the origin of a spherical coordinate system (Fig. 3.3) and by choosing our closed surface as a sphere of radius a. The electric field intensity of the point charge has been found to be

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2}\mathbf{a}_r$$

and since

$$\mathbf{D} = \epsilon_0 \mathbf{E}$$

we have, as before,

$$\mathbf{D} = \frac{Q}{4\pi r^2}\mathbf{a}_r$$

At the surface of the sphere,

$$\mathbf{D}_S = \frac{Q}{4\pi a^2}\mathbf{a}_r$$

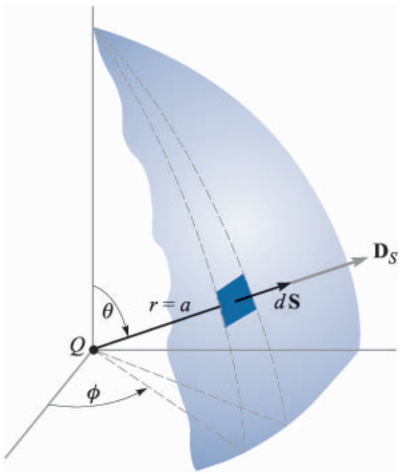


FIGURE 3.3
Application of Gauss's law to the field of a point charge Q on a spherical closed surface of radius a. The electric flux density **D** is everywhere normal to the spherical surface and has a constant magnitude at every point on it.

The differential element of area on a spherical surface is, in spherical coordinates from Chap. 1,

$$dS = r^2 \sin\theta \, d\theta \, d\phi = a^2 \sin\theta \, d\theta \, d\phi$$

or

$$d\mathbf{S} = a^2 \sin\theta \, d\theta \, d\phi$$

The integrand is

$$\mathbf{D}_S \cdot d\mathbf{S} = \frac{Q}{4\pi a^2} a^2 \sin\theta \, d\theta \, d\phi \mathbf{a}_r \cdot \mathbf{a}_r = \frac{Q}{4\pi} \sin\theta \, d\theta \, d\phi$$

leading to the closed surface integral

$$\int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \sin\theta \, d\theta \, d\phi$$

where the limits on the integrals have been chosen so that the integration is carried over the entire surface of the sphere once.[2] Integrating gives

$$\int_0^{2\pi} \frac{Q}{4\pi} (-\cos\theta)_0^{\pi} \, d\phi = \int_0^{2\pi} \frac{Q}{2\pi} \, d\phi = Q$$

and we obtain a result showing that Q coulombs of electric flux are crossing the surface, as we should since the enclosed charge is Q coulombs.

The following section contains examples of the application of Gauss's law to problems of a simple symmetrical geometry with the object of finding the elelctric field intensity.

✔ **D3.3.** Given the electric flux density, $\mathbf{D} = 0.3r^2\mathbf{a}_r \text{ nC/m}^2$ in free space: (*a*) find $\mathbf{E}$ at point $P(r = 2, \theta = 25^\circ, \phi = 90^\circ)$; (*b*) find the total charge within the sphere $r = 3$; (*c*) find the total electric flux leaving the sphere $r = 4$.

Ans. $135.5\mathbf{a}_r$ V/m; 305 nC; 965 nC

✔ **D3.4.** Calculate the total electric flux leaving the cubical surface formed by the six planes $x, y, z = \pm 5$ if the charge distribution is: (*a*) two point charges, 0.1 μC at $(1, -2, 3)$ and $\frac{1}{7}$ μC at $(-1, 2, -2)$; (*b*) a uniform line charge of π μC/m at $x = -2$, $y = 3$; (*c*) a uniform surface charge of 0.1 μC/m² on the plane $y = 3x$.

Ans. 0.243 μC; 31.4 μC; 10.54 μC

---

[2] Note that if θ and ϕ both cover the range from 0 to 2π, the spherical surface is covered twice.

3.3 APPLICATION OF GAUSS'S LAW: SOME SYMMETRICAL CHARGE DISTRIBUTIONS

Let us now consider how we may use Gauss's law,

$$Q = \oint_S \mathbf{D}_S \cdot d\mathbf{S}$$

to determine $\mathbf{D}_S$ if the charge distribution is known. This is an example of an integral equation in which the unknown quantity to be determined appears inside the integral.

The solution is easy if we are able to choose a closed surface which satisfies two conditions:

1. $\mathbf{D}_S$ is everywhere either normal or tangential to the closed surface, so that $\mathbf{D}_S \cdot d\mathbf{S}$ becomes either $D_S dS$ or zero, respectively.
2. On that portion of the closed surface for which $\mathbf{D}_S \cdot dS$ is not zero, $D_S =$ constant.

This allows us to replace the dot product with the product of the scalars D_S and dS and then to bring D_S outside the integral sign. The remaining integral is then $\int_S dS$ over that portion of the closed surface which $\mathbf{D}_S$ crosses normally, and this is simply the area of this section of that surface.

Only a knowledge of the symmetry of the problem enables us to choose such a closed surface, and this knowledge is obtained easily by remembering that the electric field intensity due to a positive point charge is directed radially outward from the point charge.

Let us again consider a point charge Q at the origin of a spherical coordinate system and decide on a suitable closed surface which will meet the two requirements listed above. The surface in question is obviously a spherical surface, centered at the origin and of any radius r. $\mathbf{D}_S$ is everywhere normal to the surface; D_S has the same value at all points on the surface.

Then we have, in order,

$$\begin{aligned} Q &= \oint_S \mathbf{D}_S \cdot d\mathbf{S} = \oint_{\text{sph}} D_S dS \\ &= D_S \oint_{\text{sph}} dS = D_S \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} r^2 \sin\theta \, d\theta \, d\phi \\ &= 4\pi r^2 D_S \end{aligned}$$

and hence $$D_S = \frac{Q}{4\pi r^2}$$

Since r may have any value and since $\mathbf{D}_S$ is directed radially outward,

$$\mathbf{D} = \frac{Q}{4\pi r^2}\mathbf{a}_r \qquad \mathbf{E} = \frac{Q}{4\pi\epsilon_0 r^2}\mathbf{a}_r$$

which agrees with the results of Chap. 2. The example is a trivial one, and the objection could be raised that we had to know that the field was symmetrical and directed radially outward before we could obtain an answer. This is true, and that leaves the inverse-square-law relationship as the only check obtained from Gauss's law. The example does, however, serve to illustrate a method which we may apply to other problems, including several to which Coulomb's law is almost incapable of supplying an answer.

Are there any other surfaces which would have satisfied our two conditions? The student should determine that such simple surfaces as a cube or a cylinder do not meet the requirements.

As a second example, let us reconsider the uniform line charge distribution ρ_L lying along the z axis and extending from $-\infty$ to $+\infty$. We must first obtain a knowledge of the symmetry of the field, and we may consider this knowledge complete when the answers to these two questions are known:

1. With which coodinates does the field vary (or of what variables is D a function)?
2. Which components of **D** are present?

These same questions were asked when we used Coulomb's law to solve this problem in Sec. 2.5. We found then that the knowledge obtained from answering them enabled us to make a much simpler integration. The problem could have been (and was) worked without any consideration of symmetry, but it was more difficult.

In using Gauss's law, however, it is not a question of using symmetry to simplify the solution, for the application of Gauss's law depends on symmetry, and *if we cannot show that symmetry exists then we cannot use Gauss's law* to obtain a solution. The two questions above now become "musts."

From our previous discussion of the uniform line charge, it is evident that only the radial component of D is present, or

$$\mathbf{D} = D_\rho \mathbf{a}_\rho$$

and this component is a function of ρ only.

$$D_\rho = f(\rho)$$

The choice of a closed surface is now simple, for a cylindrical surface is the only surface to which D_ρ is everywhere normal and it may be closed by plane surfaces normal to the z axis. A closed right circular cylindrical of radius ρ extending from $z = 0$ to $z = L$ is shown in Fig. 3.4.

We apply Gauss's law,

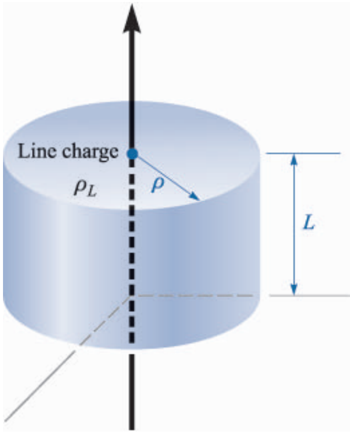


FIGURE 3.4
The gaussian surface for an infinite uniform line charge is a right circular cylinder of length L and radius ρ. **D** is constant in magnitude and everywhere perpendicular to the cylindrical surface; **D** is parallel to the end faces.

$$Q = \oint_{\text{cyl}} \mathbf{D}_S \cdot d\mathbf{S} = D_S \int_{\text{sides}} dS + 0 \int_{\text{top}} dS + 0 \int_{\text{bottom}} dS$$

$$= D_S \int_{z=0}^{L} \int_{\phi=0}^{2\pi} \rho \, d\phi \, dz = D_S 2\pi\rho L$$

and obtain
$$D_S = D_\rho = \frac{Q}{2\pi\rho L}$$

In terms of the charge density ρ_L, the total charge enclosed is

$$Q = \rho_L L$$

giving
$$D_\rho = \frac{\rho_L}{2\pi\rho}$$

or
$$E_\rho = \frac{\rho_L}{2\pi\epsilon_0\rho}$$

Comparison with Sec. 2.4, Eq. (20), shows that the correct result has been obtained and with much less work. Once the appropriate surface has been chosen, the integration usually amounts only to writing down the area of the surface at which **D** is normal.

The problem of a coaxial cable is almost identical with that of the line charge and is an example which is extremely difficult to solve from the standpoint of Coulomb's law. Suppose that we have two coaxial cylindrical conductors, the inner of radius a and the outer of radius b, each infinite in extent (Fig. 3.5). We shall assume a charge distribution of ρ_S on the outer surface of the inner conductor.

Symmetry considerations show us that only the D_ρ component is present and that it can be a function only of ρ. A right circular cylinder of length L and

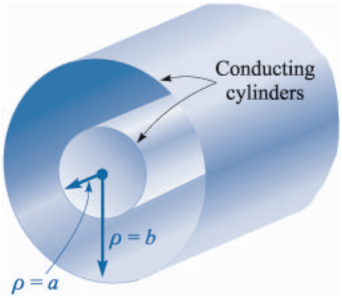


FIGURE 3.5
The two coaxial cylindrical conductors forming a coaxial cable provide an electric flux density within the cylinders, given by $D_\rho = a\rho_S/\rho$.

radius ρ, where $a < \rho < b$, is necessarily chosen as the gaussian surface, and we quickly have

$$Q = D_S 2\pi\rho L$$

The total charge on a length L of the inner conductor is

$$Q = \int_{z=0}^{L} \int_{\phi=0}^{2\pi} \rho_S a \, d\phi \, dz = 2\pi a L \rho_S$$

from which we have

$$D_S = \frac{a\rho_S}{\rho} \qquad \mathbf{D} = \frac{a\rho_S}{\rho}\mathbf{a}_\rho \qquad (a < \rho < b)$$

This result might be expressed in terms of charge per unit length, because the inner conductor has $2\pi a\rho_S$ coulombs on a meter length, and hence, letting $\rho_L = 2\pi a \rho_S$,

$$\boxed{\mathbf{D} = \frac{\rho_L}{2\pi\rho}\mathbf{a}_\rho}$$

and the solution has a form identical with that of the infinite line charge.

Since every line of electric flux starting from the charge on the inner cylinder must terminate on a negative charge on the inner surface of the outer cylinder, the total charge on that surface must be

$$Q_{\text{outer cyl}} = -2\pi a L \rho_{S,\text{inner cyl}}$$

and the surface charge on the outer cylinder is found as

$$2\pi b L \rho_{S,\text{outer cyl}} = -2\pi a L \rho_{S,\text{inner cyl}}$$

or

$$\rho_{S,\text{outer cyl}} = -\frac{a}{b}\rho_{S,\text{inner cyl}}$$

What would happen if we should use a cylinder of radius ρ, $\rho > b$, for the gaussian surface? The total charge enclosed would then be zero, for there are equal and opposite charges on each conducting cylinder. Hence

$$0 = D_S 2\pi\rho L \qquad (\rho > b)$$
$$D_S = 0 \qquad (\rho > b)$$

An identical result would be obtained for $\rho < a$. Thus the coaxial cable or capacitor has no external field (we have proved that the outer conductor is a "shield"), and there is no field within the center conductor.

Our result is also useful for a *finite* length of coaxial cable, open at both ends, provided the length L is many times greater than the radius b so that the unsymmetrical conditions at the two ends do not appreciably affect the solution. Such a device is also termed a *coaxial capacitor*. Both the coaxial cable and the coaxial capacitor will appear frequently in the work that follows.

Perhaps a numerical example can illuminate some of these results.

Example 3.2

Let us select a 50-cm length of coaxial cable having an inner radius of 1 mm and an outer radius of 4 mm. The space between conductors is assumed to be filled with air. The total charge on the inner conductor is 30 nC. We wish to know the charge density on each conductor, and the **E** and **D** fields.

Solution. We begin by finding the surface charge density on the inner cylinder,

$$\rho_{S,\text{inner cyl}} = \frac{Q_{\text{inner cyl}}}{2\pi a L} = \frac{30 \times 10^{-9}}{2\pi(10^{-3})(0.5)} = 9.55 \quad \mu\text{C/m}^2$$

The negative charge density on the inner surface of the outer cylinder is

$$\rho_{S,\text{outer cyl}} = \frac{Q_{\text{outer cyl}}}{2\pi b L} = \frac{-30 \times 10^{-9}}{2\pi(4 \times 10^{-3})(0.5)} = -2.39 \quad \mu\text{C/m}^2$$

The internal fields may therefore be calculated easily:

$$D_\rho = \frac{a\rho_S}{\rho} = \frac{10^{-3}(9.55 \times 10^{-6})}{\rho} = \frac{9.55}{\rho} \quad \text{nC/m}^2$$

and

$$E_\rho = \frac{D_\rho}{\epsilon_0} = \frac{9.55 \times 10^{-9}}{8.854 \times 10^{-12}\rho} = \frac{1079}{\rho} \quad \text{V/m}$$

Both of these expressions apply to the region where $1 < \rho < 4$ mm. For $\rho < 1$ mm or $\rho > 4$ mm, **E** and **D** are zero.

✔ **D3.5.** A point charge of 0.25 μC is located at $r = 0$, and uniform surface charge densities are located as follows: 2 mC/m² at $r = 1$ cm, and −0.6 mC/m² at $r = 1.8$ cm. Calculate **D** at: (*a*) $r = 0.5$ cm; (*b*) $r = 1.5$ cm; (*c*) $r = 2.5$ cm. (*d*) What uniform surface charge density should be established at $r = 3$ cm to cause $\mathbf{D} = 0$ at $r = 3.5$ cm?

Ans. $796\mathbf{a}_r$ μC/m²; $977\mathbf{a}_r$ μC/m²; $40.8\mathbf{a}_r$ μC/m²; −28.3 μC/m²

3.4 APPLICATION OF GAUSS'S LAW: DIFFERENTIAL VOLUME ELEMENT

We are now going to apply the methods of Guass's law to a slightly different type of problem—one which does not possess any symmetry at all. At first glance it might seem that our case is hopeless, for without symmetry a simple gaussian surface cannot be chosen such that the normal component of **D** is constant or zero everywhere on the surface. Without such a surface, the integral cannot be evaluated. There is only one way to circumvent these difficulties, and that is to choose such a very small closed surface that **D** is *almost* constant over the surface, and the small change in **D** may be adequately represented by using the first two terms of the Taylor's-series expansion for **D**. The result will become more nearly correct as the volume enclosed by the gaussian surface decreases, and we intend eventually to allow this volume to approach zero.

This example also differs from the preceding ones in that we shall not obtain the value of **D** as our answer, but instead receive some extremely valuable information about the way **D** varies in the region of our small surface. This leads directly to one of Maxwell's four equations, which are basic to all electromagnetic theory.

Let us consider any point P, shown in Fig. 3.6, located by a cartesian coordinate system. The value of **D** at the point P may be expressed in cartesian components, $\mathbf{D}_0 = D_{x0}\mathbf{a}_x + D_{y0}\mathbf{a}_y + D_{z0}\mathbf{a}_z$. We choose as our closed surface the small rectangular box, centered at P, having sides of lengths Δx, Δy, and Δz, and apply Gauss's law,

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q$$

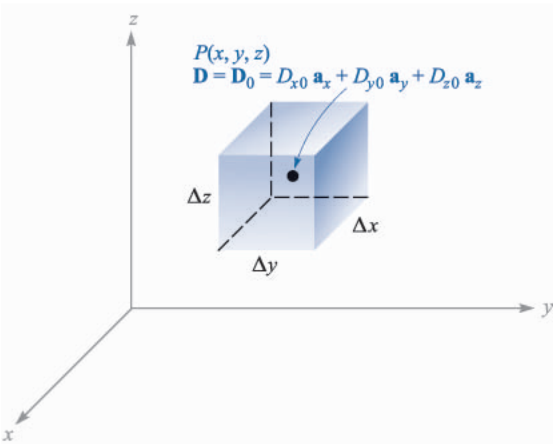


FIGURE 3.6
A differential-sized gaussian surface about the point P is used to investigate the space rate of change of **D** in the neighborhood of P.

In order to evaluate the integral over the closed surface, the integral must be broken up into six integrals, one over each face,

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_{\text{front}} + \int_{\text{back}} + \int_{\text{left}} + \int_{\text{right}} + \int_{\text{top}} + \int_{\text{bottom}}$$

Consider the first of these in detail. Since the surface element is very small, **D** is essentially constant (over *this* portion of the entire closed surface) and

$$\begin{aligned}\int_{\text{front}} &\doteq \mathbf{D}_{\text{front}} \cdot \Delta\mathbf{S}_{\text{front}} \\ &\doteq \mathbf{D}_{\text{front}} \cdot \Delta y\, \Delta z\, \mathbf{a}_x \\ &\doteq D_{x,\text{front}} \Delta y\, \Delta z\end{aligned}$$

where we have only to approximate the value of D_x at this front face. The front face is at a distance of $\Delta x/2$ from P, and hence

$$\begin{aligned}D_{x,\text{front}} &\doteq D_{x0} + \frac{\Delta x}{2} \times \text{ rate of change of } D_x \text{ with } x \\ &\doteq D_{x0} + \frac{\Delta x}{2} \frac{\partial D_x}{\partial x}\end{aligned}$$

where D_{x0} is the value of D_x at P, and where a partial derivative must be used to express the rate of change of D_x with x, since D_x in general also varies with y and z. This expression could have been obtained more formally by using the constant term and the term involving the first derivative in the Taylor's-series expansion for D_x in the neighborhood of P.

We have now

$$\int_{\text{front}} \doteq \left(D_{x0} + \frac{\Delta x}{2} \frac{\partial D_x}{\partial x} \right) \Delta y\, \Delta z$$

Consider now the integral over the back surface,

$$\begin{aligned}\int_{\text{back}} &\doteq \mathbf{D}_{\text{back}} \cdot \Delta\mathbf{S}_{\text{back}} \\ &\doteq \mathbf{D}_{\text{back}} \cdot (-\Delta y\, \Delta z\, \mathbf{a}_x) \\ &\doteq -D_{x,\text{back}} \Delta y\, \Delta z\end{aligned}$$

and

$$D_{x,\text{back}} \doteq D_{x0} - \frac{\Delta x}{2} \frac{\partial D_x}{\partial x}$$

giving

$$\int_{\text{back}} \doteq \left(-D_{x0} + \frac{\Delta x}{2} \frac{\partial D_x}{\partial x} \right) \Delta y\, \Delta z$$

If we combine these two integrals, we have

$$\int_{\text{front}} + \int_{\text{back}} \doteq \frac{\partial D_x}{\partial x} \Delta x \, \Delta y \, \Delta z$$

By exactly the same process we find that

$$\int_{\text{right}} + \int_{\text{left}} \doteq \frac{\partial D_y}{\partial y} \Delta x \, \Delta y \, \Delta z$$

and

$$\int_{\text{top}} + \int_{\text{bottom}} \doteq \frac{\partial D_z}{\partial z} \Delta x \, \Delta y \, \Delta z$$

and these results may be collected to yield

$$\oint_S \mathbf{D} \cdot d\mathbf{S} \doteq \left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \right) \Delta x \, \Delta y \, \Delta z$$

or

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q \doteq \left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \right) \Delta v \qquad (7)$$

The expression is an approximation which becomes better as Δv becomes smaller, and in the following section we shall let the volume Δv approach zero. For the moment, we have applied Gauss's law to the closed surface surrounding the volume element Δv and have as a result the approximation (7) stating that

$$\text{Charge enclosed in volume } \Delta v \doteq \left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \right) \times \text{volume } \Delta v \qquad (8)$$

Example 3.3

Find an approximate value for the total charge enclosed in an incremental volume of $10^{-9}\,\text{m}^3$ located at the origin, if $\mathbf{D} = e^{-x} \sin y \, \mathbf{a}_x - e^{-x} \cos y \, \mathbf{a}_y + 2z\mathbf{a}_z \, \text{C/m}^2$.

Solution. We first evaluate the three partial derivatives in (8):

$$\frac{\partial D_x}{\partial x} = -e^{-x} \sin y$$

$$\frac{\partial D_y}{\partial y} = e^{-x} \sin y$$

$$\frac{\partial D_z}{\partial z} = 2$$

At the origin, the first two expressions are zero, and the last is 2. Thus, we find that the charge enclosed in a small volume element there must be approximately $2\Delta v$. If Δv is $10^{-9}\,\text{m}^3$, then we have enclosed about 2 nC.

✔ **D3.6.** In free space, let $\mathbf{D} = 8xyz^4\mathbf{a}_x + 4x^2z^4\mathbf{a}_y + 16x^2yz^3$ pC/m^2. (*a*) Find the total electric flux passing through the rectangular surface $z = 2$, $0 < x < 2$, $1 < y < 3$, in the $\mathbf{a}_z$ direction. (*b*) Find $\mathbf{E}$ at $P(2, -1, 3)$. (*c*) Find an approximate value for the total charge contained in an incremental sphere located at $P(2, -1, 3)$ and having a volume of 10^{-12} m^3.

Ans. 1365 pC; $-146.4\mathbf{a}_x + 146.4\mathbf{a}_y - 195.2\mathbf{a}_z$ V/m; -2.38×10^{-21} C

3.5 DIVERGENCE

We shall now obtain an exact relationship from (7), by allowing the volume element Δv to shrink to zero. We write this equation as

$$\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right) \doteq \frac{\oint_S \mathbf{D} \cdot d\mathbf{S}}{\Delta v} = \frac{Q}{\Delta v}$$

or, as a limit

$$\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right) = \lim_{\Delta v \to 0} \frac{\oint_S \mathbf{D} \cdot d\mathbf{S}}{\Delta v} = \lim_{\Delta v \to 0} \frac{Q}{\Delta v}$$

where the approximation has been replaced by an equality. It is evident that the last term is the volume charge density ρ_v, and hence that

$$\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right) = \lim_{\Delta v \to 0} \frac{\oint_S \mathbf{D} \cdot d\mathbf{S}}{\Delta v} = \rho_v \tag{9}$$

This equation contains too much information to discuss all at once, and we shall write it as two separate equations,

$$\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right) = \lim_{\Delta v \to 0} \frac{\oint_S \mathbf{D} \cdot d\mathbf{S}}{\Delta v} \tag{10}$$

and

$$\left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}\right) = \rho_v \tag{11}$$

where we shall save (11) for consideration in the next section.

Equation (10) does not involve charge density, and the methods of the previous section could have been used on any vector $\mathbf{A}$ to find $\oint_S \mathbf{A} \cdot d\mathbf{S}$ for a small closed surface, leading to

$$\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right) = \lim_{\Delta v \to 0} \frac{\oint_S \mathbf{A} \cdot d\mathbf{S}}{\Delta v} \tag{12}$$

where $\mathbf{A}$ could represent velocity, temperature gradient, force, or any other vector field.

This operation appeared so many times in physical investigations in the last century that it received a descriptive name, *divergence*. The divergence of $\mathbf{A}$ is defined as

$$\text{Divergence of } \mathbf{A} = \text{div } \mathbf{A} = \lim_{\Delta v \to 0} \frac{\oint_S \mathbf{A} \cdot d\mathbf{S}}{\Delta v} \tag{13}$$

and is usually abbreviated div **A**. The physical interpretation of the divergence of a vector is obtained by describing carefully the operations implied by the right-hand side of (13), where we shall consider **A** as a member of the flux-density family of vectors in order to aid the physical interpretation.

> The divergence of the vector flux density **A** is the outflow of flux from a small closed surface per unit volume as the volume shrinks to zero.

The physical interpretation of divergence afforded by this statement is often useful in obtaining qualitative information about the divergence of a vector field without resorting to a mathematical investigation. For instance, let us consider the divergence of the velocity of water in a bathtub after the drain has been opened. The net outflow of water through *any* closed surface lying entirely within the water must be zero, for water is essentially incompressible and the water entering and leaving different regions of the closed surface must be equal. Hence the divergence of this velocity is zero.

If, however, we consider the velocity of the air in a tire which has just been punctured by a nail, we realize that the air is expanding as the pressure drops, and that consequently there is a net outflow from any closed surface lying within the tire. The divergence of this velocity is therefore greater than zero.

A positive divergence for any vector quantity indicates a *source* of that vector quantity at that point. Similarly, a negative divergence indicates a *sink*. Since the divergence of the water velocity above is zero, no source or sink exists.[3] The expanding air, however, produces a positive divergence of the velocity, and each interior point may be considered a source.

Writing (10) with our new term, we have

$$\text{div } \mathbf{D} = \left(\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \right) \tag{14}$$

This expression is again of a form which does not involve the charge density. It is the result of applying the definition of divergence (13) to a differential volume element in *cartesian coordinates*.

If a differential volume unit $\rho \, d\rho \, d\phi \, dz$ in cylindrical coordinates, or $r^2 \sin\theta \, dr \, d\theta \, d\phi$ in spherical coordinates, had been chosen, expressions for diver-

[3] Having chosen a differential element of volume within the water, the gradual decrease in water level with time will eventually cause the volume element to lie above the surface of the water. At the instant the surface of the water intersects the volume element, the divergence is positive and the small volume is a source. This complication is avoided above by specifying an integral point.

gence involving the components of the vector in the particular coordinate system and involving partial derivatives with respect to the variables of that system would have been obtained. These expressions are obtained in Appendix A and are given here for convenience:

$$\text{div } \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \qquad \text{(cartesian)} \tag{15}$$

$$\text{div } \mathbf{D} = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho D_\rho) + \frac{1}{\rho}\frac{\partial D_\phi}{\partial \phi} + \frac{\partial D_z}{\partial z} \qquad \text{(cylindrical)} \tag{16}$$

$$\text{div } \mathbf{D} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 D_r) + \frac{1}{r \sin\theta}\frac{\partial}{\partial \theta}(\sin\theta\, D_\theta) + \frac{1}{r \sin\theta}\frac{\partial D_\phi}{\partial \phi} \qquad \text{(spherical)} \tag{17}$$

These relationships are also shown inside the back cover for easy reference.

It should be noted that the divergence is an operation which is performed on a vector, but that the result is a scalar. We should recall that, in a somewhat similar way, the dot, or scalar, product was a multiplication of two vectors which yielded a scalar product.

For some reason it is a common mistake on meeting divergence for the first time to impart a vector quality to the operation by scattering unit vectors around in the partial derivatives. Divergence merely tells us *how much* flux is leaving a small volume on a per-unit-volume basis; no direction is associated with it.

We can illustrate the concept of divergence by continuing with the example at the end of the previous section.

Example 3.4

Find div **D** at the origin if $\mathbf{D} = e^{-x} \sin y\, \mathbf{a}_x - e^{-x} \cos y\, \mathbf{a}_y + 2z\mathbf{a}_z$.

Solution. We use (14) or (15) to obtain

$$\text{div } \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}$$
$$= -e^{-x} \sin\, y + e^{-x} \sin\, y + 2 = 2$$

The value is the constant 2, regardless of location.

If the units of **D** are C/m^2, then the units of div **D** are C/m^3. This is a volume charge density, a concept discussed in the next section.

✔ **D3.7.** In each of the following parts, find a numerical value for div **D** at the point specified: (*a*) $\mathbf{D} = (2xyz - y^2)\mathbf{a}_x + (x^2z - 2xy)\mathbf{a}_y + x^2y\mathbf{a}_z$ C/m^2 at $P_A(2, 3, -1)$; (*b*) $\mathbf{D} = 2\rho z^2 \sin^2\phi\, \mathbf{a}_\rho + \rho z^2 \sin 2\phi\, \mathbf{a}_\phi + 2\rho^2 z \sin^2\phi\, \mathbf{a}_z$ C/m^2 at $P_B(\rho = 2, \phi = 110^\circ, z = -1)$; (*c*) $\mathbf{D} = 2r\sin\theta\cos\phi\, \mathbf{a}_r + r\cos\theta\cos\phi\, \mathbf{a}_\theta - r\sin\phi\, \mathbf{a}_\phi$ at $P_C(r = 1.5, \theta = 30^\circ, \phi = 50^\circ)$.

Ans. −10.00; 9.06; 2.18

3.6 MAXWELL'S FIRST EQUATION (ELECTROSTATICS)

We now wish to consolidate the gains of the last two sections and to provide an interpretation of the divergence operation as it relates to electric flux density. The expressions developed there may be written as

$$\text{div}\,\mathbf{D} = \lim_{\Delta v \to 0} \frac{\oint_S \mathbf{D} \cdot d\mathbf{S}}{\Delta v} \tag{18}$$

$$\text{div}\,\mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} \tag{19}$$

and

$$\text{div}\,\mathbf{D} = \rho_v \tag{20}$$

The first equation is the definition of divergence, the second is the result of applying the definition to a differential volume element in cartesian coordinates, giving us an equation by which the divergence of a vector expressed in cartesian coordinates may be evaluated, and the third is merely (11) written using the new term div **D**. Equation (20) is almost an obvious result if we have achieved any familiarity at all with the concept of divergence as defined by (18), for given Gauss's law,

$$\oint_S \mathbf{A} \cdot d\mathbf{S} = Q$$

per unit volume

$$\frac{\oint_S \mathbf{A} \cdot d\mathbf{S}}{\Delta v} = \frac{Q}{\Delta v}$$

As the volume shrinks to zero,

$$\lim_{\Delta v \to 0} \frac{\oint_S \mathbf{A} \cdot d\mathbf{S}}{\Delta v} = \lim_{\Delta v \to 0} \frac{Q}{\Delta v}$$

we should see div **D** on the left and volume charge density on the right,

$$\boxed{\text{div}\,\mathbf{D} = \rho_v} \tag{20}$$

This is the first of Maxwell's four equations as they apply to electrostatics and steady magnetic fields, and it states that the electric flux per unit volume

leaving a vanishingly small volume unit is exactly equal to the volume charge density there. This equation is aptly called the *point form of Gauss's law*. Gauss's law relates the flux leaving any closed surface to the charge enclosed, and Maxwell's first equation makes an identical statement on a per-unit-volume basis for a vanishingly small volume, or at a point. Remembering that the divergence may be expressed as the sum of three partial derviatves, Maxwell's first equation is also described as the differential-equation form of Gauss's law, and conversely, Gauss's law is recognized as the integral form of Maxwell's first equation.

As a specific illustration, let us consider the divergence of **D** in the region about a point charge Q located at the origin. We have the field

$$\mathbf{D} = \frac{Q}{4\pi r^2}\mathbf{a}_r$$

and make use of (17), the expression for divergence in spherical coordinates given in the previous section:

$$\text{div } \mathbf{D} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 D_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(D_\theta \sin\theta) + \frac{1}{r\sin\theta}\frac{\partial D_\phi}{\partial\phi}$$

Since D_θ and D_ϕ are zero, we have

$$\text{div } \mathbf{D} = \frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{Q}{4\pi r^2}\right) = 0 \qquad (\text{if } r \neq 0)$$

Thus, $\rho_v = 0$ everywhere except at the origin where it is infinite.

The divergence operation is not limited to electric flux density; it can be applied to any vector field. We shall apply it to several other electromagnetic fields in the coming chapters.

✔ **D3.8.** Determine an expression for the volume charge density associated with each **D** field following: (*a*) $\mathbf{D} = \frac{4xy}{z}\mathbf{a}_x + \frac{2x^2}{z}\mathbf{a}_y - \frac{2x^2y}{z^2}\mathbf{a}_z$; (*b*) $\mathbf{D} = z\sin\phi\,\mathbf{a}_\rho + z\cos\phi\,\mathbf{a}_\phi + \rho\sin\phi\,\mathbf{a}_z$; (*c*) $\mathbf{D} = \sin\theta\sin\phi\,\mathbf{a}_r + \cos\theta\sin\phi\,\mathbf{a}_\theta + \cos\phi\,\mathbf{a}_\phi$.

Ans. $\frac{4y}{z^3}(x^2+z^2)$; 0; 0.

3.7 THE VECTOR OPERATOR ∇ AND THE DIVERGENCE THEOREM

If we remind ourselves again that divergence is an operation on a vector yielding a scalar result, just as the dot product of two vectors gives a scalar result, it seems possible that we can find something which may be dotted formally with **D** to yield the scalar

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}$$

Obviously, this cannot be accomplished by using a dot *product*; the process must be a dot *operation*.

With this in mind, we define the *del operator* ∇ as a *vector operator*,

$$\boxed{\nabla = \frac{\partial}{\partial x}\mathbf{a}_x + \frac{\partial}{\partial y}\mathbf{a}_y + \frac{\partial}{\partial z}\mathbf{a}_z} \tag{21}$$

Similar *scalar operators* appear in several methods of solving differential equations where we often let D replace d/dx, D^2 replace d^2/dx^2, and so forth.[4] We agree on defining ∇ (pronounced "del") that it shall be treated in every way as an ordinary vector with the one important exception that partial derivatives result instead of products of scalars.

Consider $\nabla \cdot \mathbf{D}$, signifying

$$\nabla \cdot \mathbf{D} = \left(\frac{\partial}{\partial x}\mathbf{a}_x + \frac{\partial}{\partial y}\mathbf{a}_y + \frac{\partial}{\partial z}\mathbf{a}_z\right) \cdot (D_x\mathbf{a}_x + D_y\mathbf{a}_y + D_z\mathbf{a}_z)$$

We first consider the dot products of the unit vectors, discarding the six zero terms and having left

$$\nabla \cdot \mathbf{D} = \frac{\partial}{\partial x}(D_x) + \frac{\partial}{\partial y}(D_y) + \frac{\partial}{\partial z}(D_z)$$

where the parentheses are now removed by operating or differentiating:

$$\nabla \cdot \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}$$

This is recognized as the divergence of $\mathbf{D}$, so that we have

$$\boxed{\text{div}\,\mathbf{D} = \nabla \cdot \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}}$$

The use of $\nabla \cdot \mathbf{D}$ is much more prevalent than that of div $\mathbf{D}$, although both usages have their advantages. Writing $\nabla \cdot \mathbf{D}$ allows us to obtain simply and quickly the correct partial derivatives, but only in cartesian coordinates, as we shall see below. On the other hand, div $\mathbf{D}$ is an excellent reminder of the physical interpretation of divergence. We shall use the operator notation $\nabla \cdot \mathbf{D}$ from now on to indicate the divergence operation.

The vector operator ∇ is used not only with divergence, but willl appear in several other very important operations later. One of these is ∇u, where u is any scalar field, and leads to

---

[4] This scalar operator D, which will not appear again, is not to be confused with the electric flux density.

$$\nabla u = \left(\frac{\partial}{\partial x}\mathbf{a}_x + \frac{\partial}{\partial y}\mathbf{a}_y + \frac{\partial}{\partial z}\mathbf{a}_z\right)u = \frac{\partial u}{\partial x}\mathbf{a}_x + \frac{\partial u}{\partial y}\mathbf{a}_y + \frac{\partial u}{\partial z}\mathbf{a}_z$$

The ∇ operator does not have a specific form in other coordinate systems. If we are considering $\mathbf{D}$ in cylindrical coordinates, then $\nabla \cdot \mathbf{D}$ still indicates the divergence of $\mathbf{D}$, or

$$\nabla \cdot \mathbf{D} = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho D_\rho) + \frac{1}{\rho}\frac{\partial D_\phi}{\partial \phi} + \frac{\partial D_z}{\partial z}$$

where this expression has been taken from Sec. 3.5. We have no form for ∇ itself to help us obtain this sum of partial derivatives. This means that ∇u, as yet unnamed but easily written above in cartesian coordinates, cannot be expressed by us at this time in cylindrical coordinates. Such an expression will be obtained when ∇u is defined in Chap. 4.

We shall close our discussion of divergence by presenting a theorem which will be needed several times in later chapters, the *divergence theorem*. This theorem applies to any vector field for which the appropriate partial derivatives exist, although it is easiest for us to develop it for the electric flux density. We have actually obtained it already and now have little more to do than point it out and name it, for starting from Gauss's law,

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q$$

and letting

$$Q = \int_{\text{vol}} \rho_v dv$$

and then replacing ρ_v by its equal,

$$\nabla \cdot \mathbf{D} = \rho_v$$

we have
$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q = \int_{\text{vol}} \rho_v dv = \int_{\text{vol}} \nabla \cdot \mathbf{D}\, dv$$

The first and last expressions constitute the divergence theorem,

$$\boxed{\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_{\text{vol}} \nabla \cdot \mathbf{D}\, dv} \qquad (22)$$

which may be stated as follows:

> The integral of the normal component of any vector field over a closed surface is equal to the integral of the divergence of this vector field throughout the volume enclosed by the closed surface.

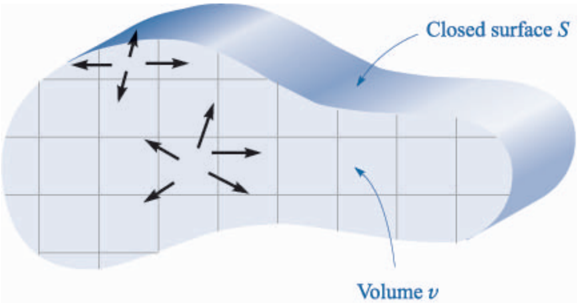


FIGURE 3.7
The divergence theorem states that the total flux crossing the closed surface is equal to the integral of the divergence of the flux density throughout the enclosed volume. The volume is shown here in cross section.

Again, we emphasize that the divergence theorem is true for any vector field, although we have obtained it specifically for the electric flux density **D**, and we shall have occasion later to apply it to several different fields. Its benefits derive from the fact that it relates a triple integration *throughout some volume* to a double integration *over the surface* of that volume. For example, it is much easier to look for leaks in a bottle full of some agitated liquid by an inspection of the surface than by calculating the velocity at every internal point.

The divergence theorem becomes obvious physically if we consider a volume v, shown in cross section in Fig. 3.7, which is surrounded by a closed surface S. Division of the volume into a number of small compartments of differential size and consideration of one cell show that the flux diverging from such a cell *enters*, or *converges* on, the adjacent cells unless the cell contains a portion of the outer surface. In summary, the divergence of the flux density throughout a volume leads, then, to the same result as determining the net flux crossing the enclosing surface.

Let us consider an example to illustrate the divergence theorem.

Example 3.5

Evaluate both sides of the divergence theorem for the field $\mathbf{D} = 2xy\mathbf{a}_x + x^2\mathbf{a}_y$ C/m^2 and the rectangular parellelepiped formed by the planes $x = 0$ and 1, $y = 0$ and 2, and $z = 0$ and 3.

Solution. Evaluating the surface integral first, we note that **D** is parallel to the surfaces at $z = 0$ and $z = 3$, so $\mathbf{D} \cdot d\mathbf{S} = 0$ there. For the remaining four surfaces we have

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_0^3 \int_0^2 (\mathbf{D})_{x=0} \cdot (-dy\, dz\, \mathbf{a}_x) + \int_0^3 \int_0^2 (\mathbf{D})_{x=1} \cdot (dy\, dz\, \mathbf{a}_x)$$
$$+ \int_0^3 \int_0^1 (\mathbf{D})_{y=0} \cdot (-dx\, dz\, \mathbf{a}_y) + \int_0^3 \int_0^1 (\mathbf{D})_{y=2} \cdot (dx\, dz\, \mathbf{a}_y)$$

$$= -\int_0^3 \int_0^2 (D_x)_{x=0} dy\, dz + \int_0^3 \int_0^2 (D_x)_{x=1} dy\, dz$$
$$- \int_0^3 \int_0^1 (D_y)_{y=0} dx\, dz + \int_0^3 \int_0^1 (D_y)_{y=2} dx\, dz$$

However, $(D_x)_{x=0} = 0$, and $(D_y)_{y=0} = (D_y)_{y=2}$, which leaves only

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_0^3 \int_0^2 (D_x)_{x=1} dy\, dz = \int_0^3 \int_0^2 2y\, dy\, dz$$
$$= \int_0^3 4\, dz = 12$$

Since
$$\nabla \cdot \mathbf{D} = \frac{\partial}{\partial x}(2xy) + \frac{\partial}{\partial y}(x^2) = 2y$$

the volume integral becomes

$$\int_{\text{vol}} \nabla \cdot \mathbf{D}\, dv = \int_0^3 \int_0^2 \int_0^1 2y\, dx\, dy\, dz = \int_0^3 \int_0^2 2y\, dy\, dz$$
$$= \int_0^3 4\, dz = 12$$

and the check is accomplished. Remembering Gauss's law, we see that we have also determined that a total charge of 12 C lies within this parallelepiped.

✔ **D3.9.** Given the field $\mathbf{D} = 6\rho \sin \frac{1}{2}\phi\, \mathbf{a}_\rho + 1.5\rho \cos \frac{1}{2}\phi\, \mathbf{a}_\phi$ C/m^2, evaluate both sides of the divergence theorem for the region bounded by $\rho = 2$, $\phi = 0$, $\phi = \pi$, $z = 0$, and $z = 5$.

Ans. 225; 225

SUGGESTED REFERENCES

1. Kraus, J. D. and D. A. Fleisch: "Electromagnetics," 5th ed., McGraw-Hill Book Company, New York, 1999. The static electric field in free space is introduced in chap. 2.
2. Plonsey, R., and R. E. Collin: "Principles and Applications of Electromagnetic Fields," McGraw-Hill Book Company, New York, 1961. The level of this text is somewhat higher than the one we are reading now, but it is an excellent text to read next. Gauss's law appears in the second chapter.
3. Plonus, M. A.: "Applied Electromagnetics," McGraw-Hill Book Company, New York, 1978. This book contains rather detailed descriptions of many practical devices that illustrate electromagnetic applications. For example, see the discussion of xerography on pp. 95–98 as an electrostatics application.
4. Skilling, H. H.: "Fundamentals of Electric Waves," 2d ed., John Wiley & Sons, Inc., New York, 1948. The operations of vector calculus are well

illustrated. Divergence is discussed on pp. 22 and 38. Chapter 1 is interesting reading.

5. Thomas, G. B., Jr., and R. L. Finney: (see Suggested References for Chap. 1). The divergence theorem is developed and illustrated from several different points of vew on pp. 976–980.

PROBLEMS

3.1 An empty metal paint can is placed on a marble table, the lid is removed, and both parts are discharged (honorably) by touching them to ground. An insulating nylon thread is glued to the center of the lid, and a penny, a nickel, and a dime are glued to the thread so that they are not touching each other. The penny is given a charge of +5 nC, and the nickel and dime are discharged. The assembly is lowered into the can so that the coins hang clear of all walls, and the lid is secured. The outside of the can is again touched momentarily to ground. The device is carefully disassembled with insulating gloves and tools. (*a*) What charges are found on each of the five metallic pieces? (*b*) If the penny had been given a charge of +5 nC, the dime a charge of −2 nC, and the nickel a charge of −1 nC, what would the final charge arrangement have been?

3.2 A point charge of 12 nC is located at the origin. Four uniform line charges are located in the $x = 0$ plane as follows: 80 nC/m at $y = -1$ and −5 m, −50 nC/m at $y = -2$ and −4 m. (*a*) Find $\mathbf{D}$ at $P(0, -3, 2)$. (*b*) How much electric flux crosses the plane $y = -3$, and in what direction? (*c*) How much electric flux leaves the surface of a sphere, 4 m in radius, centered at $C(0, -3, 0)$?

3.3 The cylindrical surface $\rho = 8$ cm contains the surface charge density, $\rho_S = 5e^{-20|z|}$ nC/m^2. (*a*) What is the total amount of charge present? (*b*) How much electric flux leaves the surface $\rho = 8$ cm, 1 cm $< z < 5$ cm, $30^\circ < \phi < 90^\circ$?

3.4 The cylindrical surfaces $\rho = 1$, 2, and 3 cm carry uniform surface charge densities of 20, −8, and 5 nC/m^2, respectively. (*a*) How much electric flux passes through the closed surface $\rho = 5$ cm, $0 < z < 1$ m? (*b*) Find $\mathbf{D}$ at $P(1\text{ cm}, 2\text{ cm}, 3\text{ cm})$.

3.5 Let $\mathbf{D} = 4xy\mathbf{a}_x + 2(x^2 + z^2)\mathbf{a}_y + 4yz\mathbf{a}_z$ C/m^2 and evaluate surface integrals to find the total charge enclosed in the rectangular parallelepiped $0 < x < 2$, $0 < y < 3$, $0 < z < 5$ m.

3.6 Two uniform line charges, each 20 nC/m, are located at $y = 1$, $z = \pm 1$ m. Find the total electric flux leaving the surface of a sphere having a radius of 2 m, if it is centered at: (*a*) $A(3, 1, 0)$; (*b*) $B(3, 2, 0)$.

3.7 Volume charge density is located in free space as $\rho_v = 2e^{-1000r}$ nC/m^3 for $0 < r < 1$ mm, and $\rho_v = 0$ elsewhere. (*a*) Find the total charge enclosed by the spherical surface $r = 1$ mm. (*b*) By using Gauss's law, calculate the value of D_r on the surface $r = 1$ mm.

3.8 Uniform line charges of 5 nC/m are located in free space at $x = 1$, $z = 1$, and at $y = 1$, $z = 0$. (*a*) Obtain an expression for **D** in cartesian coordinates at $P(0, 0, z)$. (*b*) Plot $|\mathbf{D}|$ versus z at P, $-3 < z < 10$.

3.9 A uniform volume charge density of 80 μC/m^3 is present throughout the region 8 mm $< r <$ 10 mm. Let $\rho_v = 0$ for $0 < r < 8$ mm. (*a*) Find the total charge inside the spherical surface $r = 10$ mm. (*b*) Find D_r at $r = 10$ mm. (*c*) If there is no charge for $r > 10$ mm, find D_r at $r = 20$ mm.

3.10 Let $\rho_S = 8\,\mu$C/m^2 in the region where $x = 0$ and $-4 < z < 4$ m, and let $\rho_S = 0$ elsewhere. Find **D** at $P(x, 0, z)$, where $x > 0$.

3.11 In cylindrical coordinates, let $\rho_v = 0$ for $\rho < 1$ mm, $\rho_v = 2 \sin 2000 \pi\rho$ nC/m^3 for 1 mm $< \rho <$ 1.5 mm, and $\rho_v = 0$ for $\rho > 1.5$ mm. Find **D** everywhere.

3.12 A nonuniform volume charge density, $\rho_v = 120r$ C/m^3, lies within the spherical surface $r = 1$ m, and $\rho_v = 0$ elsewhere. (*a*) Find D_r everywhere. (*b*) What surface charge density ρ_{S2} should be on the surface $r = 2$ so that $D_{r|r=2^-} = 2D_{r|r=2^+}$? (*c*) Make a sketch of D_r vs r for $0\ r < 5$ with both distributions present.

3.13 Spherical surfaces at $r = 2$, 4, and 6 m carry uniform surface charge densities of 20 nC/m^2, -4 nC/m^2, and ρ_{S0}, respectively. (*a*) Find **D** at $r = 1$, 3, and 5 m. (*b*) Determine ρ_{S0} such that $\mathbf{D} = 0$ at $r = 7$ m.

3.14 If $\rho_v = 5$ nC/m^3 for $0 < \rho < 1$ mm and no other charges are present: (*a*) find D_ρ for $\rho < 1$ mm; (*b*) find D_ρ for $\rho > 1$ mm. (*c*) What line charge ρ_L at $\rho = 0$ would give the same result for part *b*?

3.15 Volume charge density is located as follows: $\rho_v = 0$ for $\rho < 1$ mm and for $\rho > 2$ mm, $\rho_v = 4\rho\,\mu$C/m^3 for $1 < \rho < 2$ mm. (*a*) Calculate the total charge in the region, $0 < \rho < \rho_1$, $0 < z < L$, where $1 < \rho_1 < 2$ mm. (*b*) Use Gauss's law to determine D_ρ at $\rho = \rho_1$. (*c*) Evaluate D_ρ at $\rho = 0.8$ mm, 1.6 mm, and 2.4 mm.

3.16 Given the electric flux density, $\mathbf{D} = 2xy\mathbf{a}_x + x^2\mathbf{a}_y + 6z^3\mathbf{a}_z$ C/m^2: (*a*) use Gauss's law to evaluate the total charge enclosed in the volume $0 < x, y, z < a$; (*b*) use Eq. (8) to find an approximate value for the above charge. Evaluate the derivatives at $P(a/2, a/2, a/2)$. (*c*) Show that the results of parts *a* and *b* agree in the limit as $a \to 0$.

3.17 A cube is defined by $1 < x, y, z < 1.2$. If $\mathbf{D} = 2x^2y\mathbf{a}_x + 3x^2y^2\mathbf{a}_y$ C/m^2: (*a*) apply Gauss's law to find the total flux leaving the closed surface of the cube; (*b*) evaluate $\dfrac{\partial D_x}{\partial x} + \dfrac{\partial D_y}{\partial y} + \dfrac{\partial D_z}{\partial z}$ at the center of the cube. (*c*) Estimate the total charge enclosed within the cube by using Eq. (8).

3.18 Let a vector field be given by $\mathbf{G} = 5x^4y^4z^4\mathbf{a}_y$. Evaluate both sides of Eq. (8) for this **G** field and the volume defined by $x = 3$ and 3.1, $y = 1$ and 1.1, and $z = 2$ and 2.1. Evaluate the partial derivatives at the center of the volume.

3.19 A spherical surface of radius 3 mm is centered at $P(4, 1, 5)$ in free space. Let $\mathbf{D} = x\mathbf{a}_x$ C/m^2. Use the results of Sec. 3.4 to estimate the net electric flux leaving the spherical surface.

3.20 A cube of volume a^3 has its faces parallel to the cartesian coordinate surfaces. It is centered at $P(3, -2, 4)$. Given the field $\mathbf{D} = 2x^3\mathbf{a}_x$ C/m^2: (*a*) calculate div $\mathbf{D}$ at P; (*b*) evaluate the fraction in the rightmost side of Eq. (13) for $a = 1$ m, 0.1 m, and 1 mm.

3.21 Calculate the divergence of $\mathbf{D}$ at the point specified if $\mathbf{D} =$: (*a*) $\frac{1}{z^2}[10xyz\mathbf{a}_x + 5x^2z\mathbf{a}_y + (2z^3 - 5x^2y)]$ at $P(-2, 3, 5)$; (*b*) $5z^2\mathbf{a}_\rho + 10\rho z\mathbf{a}_z$ at $P(3, -45°, 5)$; (*c*) $2r\sin\theta\sin\phi\mathbf{a}_r + r\cos\theta\sin\phi\mathbf{a}_\theta + r\cos\phi\mathbf{a}_\phi$ at $P(3, 45°, -45°)$.

3.22 Let $\mathbf{D} = 8\rho\sin\phi\mathbf{a}_\rho + 4\rho\cos\phi\mathbf{a}_\phi$ C/m^2. (*a*) Find div $\mathbf{D}$. (*b*) Find the volume charge density at $P(2.6, 38°, -6.1)$. (*c*) How much charge is located inside the region defined by $0 < \rho < 1.8$, $20° < \phi < 70°$, $2.4 < z < 3.1$?

3.23 (*a*) A point charge Q lies at the origin. Show that div $\mathbf{D} = 0$ everywhere except at the origin. (*b*) Replace the point charge with a uniform volume charge density ρ_{v0} for $0 \le r \le a$. Relate ρ_{v0} to Q and a so that the total charge is the same. Find div $\mathbf{D}$ everywhere.

3.24 Inside the cylindrical shell, $3 < \rho < 4$ m, the electric flux density is given as $5(\rho - 3)^3\mathbf{a}_\rho$ C/m^2. (*a*) What is the volume charge density at $\rho = 4$ m? (*b*) What is the electric flux density at $\rho = 4$ m? (*c*) How much electric flux leaves the closed surface: $3 < \rho < 4$, $0 < \phi < 2\pi$, $-2.5 < z < 2.5$? (*d*) How much charge is contained within the volume $3 < \rho < 4$, $0 < \phi < 2\pi$, $-2.5 < z < 2.5$?

3.25 Within the spherical shell, $3 < r < 4$ m, the electric flux density is given as $\mathbf{D} = 5(r - 3)^3\mathbf{a}_r$ C/m^2. (*a*) What is the volume charge density at $r = 4$? (*b*) What is the electric flux density at $r = 4$? (*c*) How much electric flux leaves the sphere $r = 4$? (*d*) How much charge is contained within the sphere $r = 4$?

3.26 Given the field, $\mathbf{D} = \frac{5\sin\theta\cos\phi}{r}\mathbf{a}_r$ C/m^2, find: (*a*) the volume charge density; (*b*) the total charge contained in the region $r < 2$ m; (*c*) the value of $\mathbf{D}$ at the surface $r = 2$; (*d*) the total electric flux leaving the surface $r = 2$.

3.27 Let $\mathbf{D} = 5r^2\mathbf{a}_r$ mC/m^2 for $r < 0.08$ m, and $\mathbf{D} = 0.1\mathbf{a}_r/r^2$ C/m^2 for $r > 0.08$ m. (*a*) Find ρ_v for $r = 0.06$ m. (*b*) Find ρ_v for $r = 0.1$ m. (*c*) What surface charge density could be located at $r = 0.08$ m to cause $\mathbf{D} = 0$ for $r > 0.08$ m?

3.28 The electric flux density is given as $\mathbf{D} = 20\rho^3\mathbf{a}_\rho$ C/m^2 for $\rho < 100\ \mu$m, and $k\mathbf{a}_\rho/\rho$ for $\rho > 100\ \mu$m. (*a*) Find k so that $\mathbf{D}$ is continuous at $\rho = 100\ \mu$m. (*b*) Find and sketch ρ_v as a function of ρ.

3.29 In the region of free space that includes the volume, $2 < x, y, z < 3$, $\mathbf{D} = \frac{2}{z^2}(yz\mathbf{a}_x + xz\mathbf{a}_y - 2xy\mathbf{a}_z)$ C/m^2. (*a*) Evaluate the volume-integral

side of the divergence theorem for the volume defined by $2 < x, y, z < 3$. (*b*) Evaluate the surface-integral side for the corresponding closed surface.

3.30 If $\mathbf{D} = 15\rho^2 \sin 2\phi \mathbf{a}_\rho + 10\rho^2 \cos 2\phi \mathbf{a}_\phi$ C/m^2, evaluate both sides of the divergence theorem for the region: $1 < \rho < 2$ m, $1 < \phi < 2$ rad, $1 < z < 2$ m.

3.31 Given the flux density, $\mathbf{D} = \frac{16}{r} \cos 2\theta \mathbf{a}_\theta$ C/m^2, use two different methods to find the total charge within the region $1 < r < 2$ m, $1 < \theta < 2$ rad, $1 < \phi < 2$ rad.

3.32 If $\mathbf{D} = 2r\mathbf{a}_r$ C/m^2, find the total electric flux leaving the surface of the cube, $0 \leq x, y, z \leq 0.4$.

CHAPTER

4

ENERGY AND POTENTIAL

In the previous two chapters we became acquainted with Coulomb's law and its use in finding the electric field about several simple distributions of charge, and also with Gauss's law and its application in determining the field about some symmetrical charge arrangements. The use of Gauss's law was invariably easier for these highly symmetrical distributions, because the problem of integration always disappeared when the proper closed surface was chosen.

However, if we had attempted to find a slightly more complicated field, such as that of two unlike point charges separated by a small distance, we would have found it impossible to choose a suitable gaussian surface and obtain an answer. Coulomb's law, however, is more powerful and enables us to solve problems for which Gauss's law is not applicable. The application of Coulomb's law is laborious, detailed, and often quite complex, the reason for this being precisely the fact that the electric field intensity, a vector field, must be found directly from the charge distribution. Three different integrations are needed in general, one for each component, and the resolution of the vector into components usually adds to the complexity of the integrals.

Certainly it would be desirable if we could find some as yet undefined scalar function with a single integration and then determine the electric field from this scalar by some simple straightforward procedure, such as differentiation.

This scalar function does exist and is known as the *potential* or *potential field*. We shall find that it has a very real physical interpretation and is more familiar to most of us than is the electric field which it will be used to find.

We should expect, then, to be equipped soon with a third method of finding electric fields—a single scalar integration, although not always as simple as we might wish, followed by a pleasant differentiation.

The remaining difficult portion of the task, the integration, we intend to remove in Chap. 7.

4.1 ENERGY EXPENDED IN MOVING A POINT CHARGE IN AN ELECTRIC FIELD

The electric field intensity was defined as the force on a unit test charge at that point at which we wish to find the value of this vector field. If we attempt to move the test charge against the electric field, we have to exert a force equal and opposite to that exerted by the field, and this requires us to expend energy, or do work. If we wish to move the charge in the direction of the field, our energy expenditure turns out to be negative; we do not do the work, the field does.

Suppose we wish to move a charge Q a distance $d\mathbf{L}$ in an electric field $\mathbf{E}$. The force on Q due to the electric field is

$$\boxed{\mathbf{F}_E = Q\mathbf{E}} \tag{1}$$

where the subscript reminds us that this force is due to the field. The component of this force in the direction $d\mathbf{L}$ which we must overcome is

$$F_{EL} = \mathbf{F} \cdot \mathbf{a}_L = Q\mathbf{E} \cdot \mathbf{a}_L$$

where $\mathbf{a}_L =$ a unit vector in the direction of $d\mathbf{L}$.

The force which we must apply is equal and opposite to the force due to the field,

$$F_{\text{appl}} = -Q\mathbf{E} \cdot \mathbf{a}_L$$

and our expenditure of energy is the product of the force and distance. That is,

$$\text{Differential work done by external source moving } Q$$
$$= -Q\mathbf{E} \cdot \mathbf{a}_L dL = -Q\mathbf{E} \cdot d\mathbf{L}$$

or

$$\boxed{dW = -Q\mathbf{E} \cdot d\mathbf{L}} \tag{2}$$

where we have replaced $\mathbf{a}_L dL$ by the simpler expression $d\mathbf{L}$.

This differential amount of work required may be zero under several conditions determined easily from (2). There are the trivial conditions for which $\mathbf{E}$, Q, or $d\mathbf{L}$ is zero, and a much more important case in which $\mathbf{E}$ and $d\mathbf{L}$ are

perpendicular. Here the charge is moved always in a direction at right angles to the electric field. We can draw on a good analogy between the electric field and the gravitational field, where, again, energy must be expended to move against the field. Sliding a mass around with constant velocity on a frictionless surface is an effortless process if the mass is moved along a constant elevation contour; positive or negative work must be done in moving it to a higher or lower elevation, respectively.

Returning to the charge in the electric field, the work required to move the charge a finite distance must be determined by integrating,

$$W = -Q\int_{\text{init}}^{\text{final}} \mathbf{E}\cdot d\mathbf{L} \tag{3}$$

where the path must be specified before the integral can be evaluated. The charge is assumed to be at rest at both its initial and final positions.

This definite integral is basic to field theory, and we shall devote the following section to its interpretation and evaluation.

✔ **D4.1.** Given the electric field $\mathbf{E} = \frac{1}{z^2}(8xyz\mathbf{a}_x + 4x^2z\mathbf{a}_y - 4x^2y\mathbf{a}_z)$ V/m, find the differential amount of work done in moving a 6-nC charge a distance of 2 μm, starting at $P(2, -2, 3)$ and proceeding in the direction $\mathbf{a}_L =$: (*a*) $-\frac{6}{7}\mathbf{a}_x + \frac{3}{7}\mathbf{a}_y + \frac{2}{7}\mathbf{a}_z$; (*b*) $\frac{6}{7}\mathbf{a}_x - \frac{3}{7}\mathbf{a}_y - \frac{2}{7}\mathbf{a}_z$; (*c*) $\frac{3}{7} + \frac{6}{7}\mathbf{a}_y$.

Ans. −149.3; 149.3; 0 fJ

4.2 THE LINE INTEGRAL

The integral expression for the work done in moving a point charge Q from one position to another, Eq. (3), is an example of a line integral, which in vector-analysis notation always takes the form of the integral along some prescribed path of the dot product of a vector field and a differential vector path length $d\mathbf{L}$. Without using vector analysis we should have to write

$$W = -Q\int_{\text{init}}^{\text{final}} E_L\, dL$$

where E_L = component of $\mathbf{E}$ along $d\mathbf{L}$.

A line integral is like many other integrals which appear in advanced analysis, including the surface integral appearing in Gauss's law, in that it is essentially descriptive. We like to look at it much more than we like to work it out. It tells us to choose a path, break it up into a large number of very small segments, multiply the component of the field along each segment by the length of the segment, and then add the results for all the segments. This is a summation, of course, and the integral is obtained exactly only when the number of segments becomes infinite.

This procedure is indicated in Fig. 4.1, where a path has been chosen from an initial position B to a final position[1] A and a *uniform electric field* selected for simplicity. The path is divided into six segments, $\Delta\mathbf{L}_1, \Delta\mathbf{L}_2, \ldots, \Delta\mathbf{L}_6$, and the components of $\mathbf{E}$ along each segment denoted by $E_{L1}, E_{L2}, \ldots, E_{L6}$. The work involved in moving a charge Q from B to A is then approximately

$$W = -Q(E_{L1}\Delta L_1 + E_{L2}\Delta L_2 + \ldots + E_{L6}\Delta L_6)$$

or, using vector notation,

$$W = -Q(\mathbf{E}_1 \cdot \Delta\mathbf{L}_1 + \mathbf{E}_2 \cdot \Delta\mathbf{L}_2 + \ldots + \mathbf{E}_6 \cdot \Delta\mathbf{L}_6)$$

and since we have assumed a uniform field,

$$\mathbf{E}_1 = \mathbf{E}_2 = \ldots = \mathbf{E}_6$$

$$W = -Q\mathbf{E} \cdot (\Delta\mathbf{L}_1 + \Delta\mathbf{L}_2 + \ldots + \Delta\mathbf{L}_6)$$

What is this sum of vector segments in the parentheses above? Vectors add by the parallelogram law, and the sum is just the vector directed from the initial point B to the final point A, $\mathbf{L}_{BA}$. Therefore

$$W = -Q\mathbf{E} \cdot \mathbf{L}_{BA} \qquad \text{(uniform } \mathbf{E}\text{)} \tag{4}$$

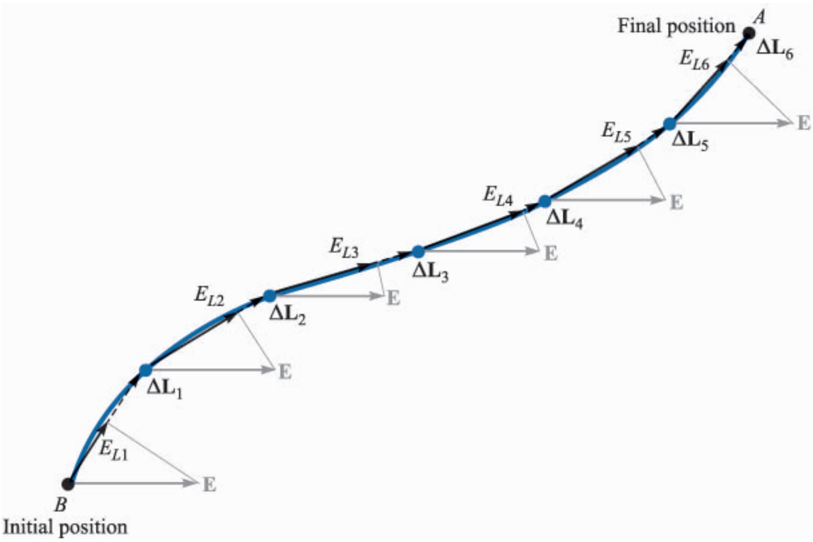


FIGURE 4.1
A graphical interpretation of a line integral in a uniform field. The line integral of **E** between points B and A is independent of the path selected, even in a nonuniform field; this result is not, in general, true for time-varying fields.

[1] The final position is given the designation A to correspond with the convention for potential difference, as discussed in the following section.

Remembering the summation interpretation of the line integral, this result for the uniform field can be obtained rapidly now from the integral expression

$$W = -Q \int_B^A \mathbf{E} \cdot d\mathbf{L} \tag{5}$$

as applied to a uniform field

$$W = -Q\mathbf{E} \cdot \int_B^A d\mathbf{L}$$

where the last integral becomes $\mathbf{L}_{BA}$ and

$$W = -Q\mathbf{E} \cdot \mathbf{L}_{BA} \qquad \text{(uniform } \mathbf{E}\text{)}$$

For this special case of a uniform electric field intensity, we should note that the work involved in moving the charge depends only on Q, $\mathbf{E}$, and $\mathbf{L}_{BA}$, a vector drawn from the initial to the final point of the path chosen. It does not depend on the particular path we have selected along which to carry the charge. We may proceed from B to A on a straight line or via the Old Chisholm Trail; the answer is the same. We shall show in Sec. 4.5 that an identical statement may be made for any nonuniform (static) $\mathbf{E}$ field.

Let us use several examples to illustrate the mechanics of setting up the line integral appearing in (5).

Example 4.1

We are given the nonuniform field

$$\mathbf{E} = y\mathbf{a}_x + x\mathbf{a}_y + 2\mathbf{a}_z$$

and we are asked to determine the work expended in carrying 2 C from $B(1, 0, 1)$ to $A(0.8, 0.6, 1)$ along the shorter arc of the circle

$$x^2 + y^2 = 1 \quad z = 1$$

Solution. We use $W = -Q \int_B^A \mathbf{E} \cdot d\mathbf{L}$, where $\mathbf{E}$ is not necessarily constant. Working in cartesian coordinates, the differential path $d\mathbf{L}$ is $dx\mathbf{a}_x + dy\mathbf{a}_y + dz\mathbf{a}_z$, and the integral becomes

$$\begin{aligned} W &= -Q \int_B^A \mathbf{E} \cdot d\mathbf{L} \\ &= -2 \int_B^A (y\mathbf{a}_x + x\mathbf{a}_y + 2\mathbf{a}_z) \cdot (dx\,\mathbf{a}_x + dy\,\mathbf{a}_y + dz\,\mathbf{a}_z) \\ &= -2 \int_1^{0.8} y\,dx - 2 \int_0^{0.6} x\,dy - 4 \int_1^1 dz \end{aligned}$$

where the limits on the integrals have been chosen to agree with the initial and final values of the appropriate variable of integration. Using the equation of the circular path (and selecting the sign of the radical which is correct for the quadrant involved), we have

$$W = -2\int_{1}^{0.8} \sqrt{1-x^2}dx - 2\int_{0}^{0.6} \sqrt{1-y^2}dy - 0$$
$$= -\left[x\sqrt{1-x^2} + \sin^{-1} x\right]_1^{0.8} - \left[y\sqrt{1-y^2} + \sin^{-1} y\right]_0^{0.6}$$
$$= -(0.48 + 0.927 - 0 - 1.571) - (0.48 + 0.644 - 0 - 0)$$
$$= -0.96\,\text{J}$$

Example 4.2

Again find the work required to carry 2 C from B to A in the same field, but this time use the straight-line path from B to A.

Solution. We start by determining the equations of the straight line. Any two of the following three equations for planes passing through the line are sufficient to define the line:

$$y - y_B = \frac{y_A - y_B}{x_A - x_B}(x - x_B)$$
$$z - z_B = \frac{z_A - z_B}{y_A - y_B}(y - y_B)$$
$$x - x_B = \frac{x_A - x_B}{z_A - z_B}(z - z_B)$$

From the first equation above we have

$$y = -3(x - 1)$$

and from the second we obtain

$$z = 1$$

Thus,

$$W = -2\int_{1}^{0.8} y\,dx - 2\int_{0}^{0.6} x\,dy - 4\int_{1}^{1} dz$$
$$= 6\int_{1}^{0.8} (x-1)dx - 2\int_{0}^{0.6} \left(1 - \frac{y}{3}\right) dy$$
$$= -0.96\,\text{J}$$

This is the same answer we found using the circular path between the same two points, and it again demonstrates the statement (unproved) that the work done is independent of the path taken in any electrostatic field.

It should be noted that the equations of the straight line show that $dy = -3\,dx$ and $dx = -3\,dy$. These substitutions may be made in the first two integrals above, along with a change in limits, and the answer may be obtained by evaluating the new integrals. This method is often simpler if the integrand is a function of only one variable.

Note that the expressions for $d\mathbf{L}$ in our three coordinate systems utilize the differential lengths obtained in the first chapter (cartesian in Sec. 1.3, cylindrical in Sec. 1.8, and spherical in Sec. 1.9):

$$d\mathbf{L} = dx\,\mathbf{a}_x + dy\,\mathbf{a}_y + dz\,\mathbf{a}_z \qquad \text{(cartesian)} \qquad (6)$$

$$d\mathbf{L} = d\rho\,\mathbf{a}\rho + \rho\,d\phi\,\mathbf{a}_\phi + dz\,\mathbf{a}_z \qquad \text{(cylindrical)} \qquad (7)$$

$$d\mathbf{L} = dr\,\mathbf{a}_r + r\,d\theta\,\mathbf{a}_\theta + r\sin\theta\,d\phi\,\mathbf{a}_\phi \qquad \text{(spherical)} \qquad (8)$$

The interrelationships among the several variables in each expression are determined from the specific equations for the path.

As a final example illustrating the evaluation of the line integral, let us investigate several paths which we might take near an infinite line charge. The field has been obtained several times and is entirely in the radial direction,

$$\mathbf{E} = E_\rho \mathbf{a}_\rho = \frac{\rho_L}{2\pi\epsilon_0\rho}\mathbf{a}_\rho$$

Let us first find the work done in carrying the positive charge Q about a circular path of radius ρ_b centered at the line charge, as illustrated in Fig. 4.2*a*. Without lifting a pencil, we see that the work must be nil, for the path is always perpendicular to the electric field intensity, or the force on the charge is always exerted at right angles to the direction in which we are moving it. For practice, however, let us set up the integral and obtain the answer.

The differential element $d\mathbf{L}$ is chosen in cylindrical coordinates, and the circular path selected demands that $d\rho$ and dz be zero, so $d\mathbf{L} = \rho_1 d\phi\,\mathbf{a}_\phi$. The work is then

$$W = -Q\int_{\text{init}}^{\text{final}} \frac{\rho_L}{2\pi\epsilon_0\rho_1}\mathbf{a}_\rho \cdot \rho_1\,d\phi\,\mathbf{a}_\phi$$

$$= -Q\int_0^{2\pi} \frac{\rho_L}{2\pi\epsilon_0}\,d\,\phi\,\mathbf{a}_\rho \cdot \mathbf{a}_\phi = 0$$

Let us now carry the charge from $\rho = a$ to $\rho = b$ along a radial path (Fig. 4.2*b*). Here $d\mathbf{L} = d\rho\,\mathbf{a}_\rho$ and

$$W = -Q\int_{\text{init}}^{\text{final}} \frac{\rho_L}{2\pi\epsilon_0\rho}\mathbf{a}_\rho \cdot d\rho\,\mathbf{a}_\rho = -Q\int_a^b \frac{\rho_L}{2\pi\epsilon_0}\frac{d\,\rho}{\rho}$$

or

$$W = -\frac{Q\rho_L}{2\pi\epsilon_0}\ln\frac{b}{a}$$

Since b is larger than a, $\ln(b/a)$ is positive, and we see that the work done is negative, indicating that the external source that is moving the charge receives energy.

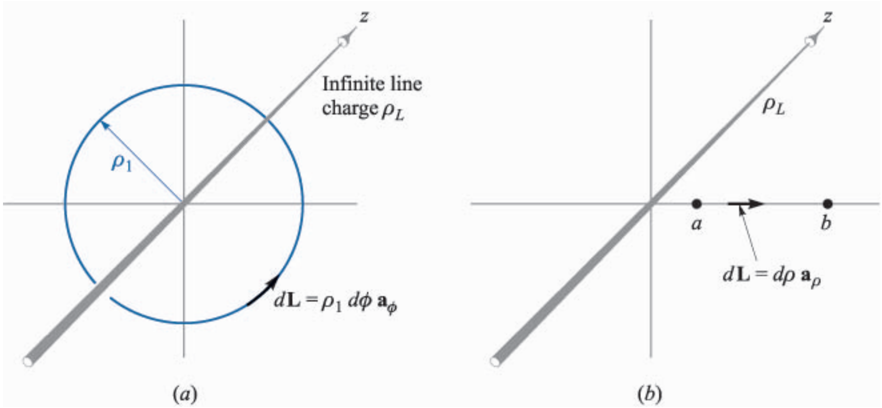


FIGURE 4.2
(*a*) A circular path and (*b*) a radial path along which a charge of Q is carried in the field of an infinite line charge. No work is expected in the former case.

One of the pitfalls in evaluating line integrals is a tendency to use too many minus signs when a charge is moved in the direction of a *decreasing* coordinate value. This is taken care of completely by the limits on the integral, and no misguided attempt should be made to change the sign of $d\mathbf{L}$. Suppose we carry Q from b to a (Fig. 4.2*b*). We still have $d\mathbf{L} = d\rho\, \mathbf{a}_\rho$ and show the different direction by recognizing $\rho = b$ as the initial point and $\rho = a$ as the final point,

$$W = -Q \int_b^a \frac{\rho_L}{2\pi\epsilon_0} \frac{d\rho}{\rho} = \frac{Q\rho_L}{2\pi\epsilon_0} \ln \frac{b}{a}$$

This is the negative of the previous answer and is obviously correct.

✔ **D4.2.** Calculate the work done in moving a 4-C charge from $B(1, 0, 0)$ to $A(0, 2, 0)$ along the path $y = 2 - 2x$, $z = 0$ in the field $\mathbf{E} =$: (*a*) $5\mathbf{a}_x$ V/m; (*b*) $5x\mathbf{a}_x$ V/m; (*c*) $5x\mathbf{a}_x + 5y\mathbf{a}_y$ V/m.

Ans. 20 J; 10 J; −30 J

✔ **D4.3.** We shall see later that a time-varying $\mathbf{E}$ field need not be conservative. (If it is not conservative, the work expressed by Eq. (3) may be a function of the path used.) Let $\mathbf{E} = y\mathbf{a}_x$ V/m at a certain instant of time, and calculate the work required to move a 3-C charge from (1, 3, 5) to (2, 0, 3) along the straight line segments joining: (*a*) (1, 3, 5) to (2, 3, 5) to (2, 0, 5) to (2, 0, 3); (*b*) (1, 3, 5) to (1, 3, 3) to (1, 0, 3) to (2, 0, 3).

Ans. −9 J; 0

4.3 DEFINITION OF POTENTIAL DIFFERENCE AND POTENTIAL

We are now ready to define a new concept from the expression for the work done by an external source in moving a charge Q from one point to another in an electric field $\mathbf{E}$,

$$W = -Q \int_{\text{init}}^{\text{final}} \mathbf{E} \cdot d\mathbf{L}$$

In much the same way as we defined the electric field intensity as the force on a *unit* test charge, we now define *potential difference* V as the work done (by an external source) in moving a *unit* positive charge from one point to another in an electric field,

$$\text{Potential difference} = V = -\int_{\text{init}}^{\text{final}} \mathbf{E} \cdot d\mathbf{L} \tag{9}$$

We shall have to agree on the direction of movement, as implied by our language, and we do this by stating that V_{AB} signifies the potential difference between points A and B and is the work done in moving the unit charge from B (last named) to A (first named). Thus, in determining V_{AB}, B is the initial point and A is the final point. The reason for this somewhat peculiar definition will become clearer shortly, when it is seen that the initial point B is often taken at infinity, whereas the final point A represents the fixed position of the charge; point A is thus inherently more significant.

Potential difference is measured in joules per coulomb, for which the *volt* is defined as a more common unit, abbreviated as V. Hence the potential difference between points A and B is

$$V_{AB} = -\int_{B}^{A} \mathbf{E} \cdot d\mathbf{L} \quad \text{V} \tag{10}$$

and V_{AB} is positive if work is done in carrying the positive charge from B to A.

From the line-charge example of the last section we found that the work done in taking a charge Q from $\rho = b$ to $\rho = a$ was

$$W = \frac{Q\rho_L}{2\pi\epsilon_0} \ln \frac{b}{a}$$

Thus, the potential difference between points at $\rho = a$ and $\rho = b$ is

$$V_{ab} = \frac{W}{Q} = \frac{\rho_L}{2\pi\epsilon_0} \ln \frac{b}{a} \tag{11}$$

We can try out this definition by finding the potential difference between points A and B at radial distances r_A and r_B from a point charge Q. Choosing an origin at Q,

$$\mathbf{E} = E_r \mathbf{a}_r = \frac{Q}{4\pi\epsilon_0 r^2} \mathbf{a}_r$$

and $$d\mathbf{L} = dr\, \mathbf{a}_r$$

we have $$V_{AB} = -\int_B^A \mathbf{E} \cdot d\mathbf{L} = -\int_{r_B}^{r_A} \frac{Q}{4\pi\epsilon_0 r^2} dr = \frac{Q}{4\pi\epsilon_0}\left(\frac{1}{r_A} - \frac{1}{r_B}\right) \tag{12}$$

If $r_B > r_A$, the potential difference V_{AB} is positive, indicating that energy is expended by the external source in bringing the positive charge from r_B to r_A. This agrees with the physical picture showing the two like charges repelling each other.

It is often convenient to speak of the *potential*, or *absolute potential*, of a point, rather than the potential difference between two points, but this means only that we agree to measure every potential difference with respect to a specified reference point which we consider to have zero potential. Common agreement must be reached on the zero reference before a statement of the potential has any significance. A person having one hand on the deflection plates of a cathode-ray tube which are "at a potential of 50 V" and the other hand on the cathode terminal would probably be too shaken up to understand that the cathode is not the zero reference, but that all potentials in that circuit are customarily measured with respect to the metallic shield about the tube. The cathode may be several thousands of volts negative with respect to the shield.

Perhaps the most universal zero reference point in experimental or physical potential measurements is "ground," by which we mean the potential of the surface region of the earth itself. Theoretically, we usually represent this surface by an infinite plane at zero potential, although some large-scale problems, such as those involving propagation across the Atlantic Ocean, require a spherical surface at zero potential.

Another widely used reference "point" is infinity. This usually appears in theoretical problems approximating a physical situation in which the earth is relatively far removed from the region in which we are interested, such as the static field near the wing tip of an airplane that has acquired a charge in flying through a thunderhead, or the field inside an atom. Working with the *gravitational* potential field on earth, the zero reference is normally taken at sea level; for an interplanetary mission, however, the zero reference is more conveniently selected at infinity.

A cylindrical surface of some definite radius may occasionally be used as a zero reference when cylindrical symmetry is present and infinity proves inconvenient. In a coaxial cable the outer conductor is selected as the zero reference for potential. And, of course, there are numerous special problems, such as those for which a two-sheeted hyperboloid or an oblate spheroid must be selected as the zero-potential reference, but these need not concern us immediately.

If the potential at point A is V_A and that at B is V_B, then

$$\boxed{V_{AB} = V_A - V_B} \tag{13}$$

where we necessarily agree that V_A and V_B shall have the same zero reference point.

✔ **D4.4.** An electric field is expressed in cartesian coordinates by $\mathbf{E} = 6x^2\mathbf{a}_x + 6y\mathbf{a}_y + 4\mathbf{a}_z$ V/m. Find: (*a*) V_{MN} if points M and N are specified by $M(2, 6, -1)$ and $N(-3, -3, 2)$; (*b*) V_M if $V = 0$ at $Q(4, -2, -35)$; (*c*) V_N if $V = 2$ at $P(1, 2, -4)$.

Ans. -139.0 V; -120.0 V; 19.00 V

4.4 THE POTENTIAL FIELD OF A POINT CHARGE

In the previous section we found an expression (12) for the potential difference between two points located at $r = r_A$ and $r = r_B$ in the field of a point charge Q placed at the origin,

$$V_{AB} = \frac{Q}{4\pi\epsilon_0}\left(\frac{1}{r_A} - \frac{1}{r_B}\right) = V_A - V_B \tag{14}$$

It was assumed that the two points lay on the same radial line or had the same θ and ϕ coordinate values, allowing us to set up a simple path on this radial line along which to carry our positive charge. We now should ask whether different θ and ϕ coordinate values for the initial and final position will affect our answer and whether we could choose more complicated paths between the two points without changing the results. Let us answer both questions at once by choosing two general points A and B (Fig. 4.3) at radial distances of r_A and r_B, and any values for the other coordinates.

The differential path length $d\mathbf{L}$ has r, θ, and ϕ components, and the electric field has only a radial component. Taking the dot product then leaves us only

$$V_{AB} = -\int_{r_B}^{r_A} E_r dr = -\int_{r_B}^{r_A} \frac{Q}{4\pi\epsilon_0 r^2} dr = \frac{Q}{4\pi\epsilon_0}\left(\frac{1}{r_A} - \frac{1}{r_B}\right)$$

We obtain the same answer and see, therefore, that the potential difference between two points in the field of a point charge depends only on the distance of

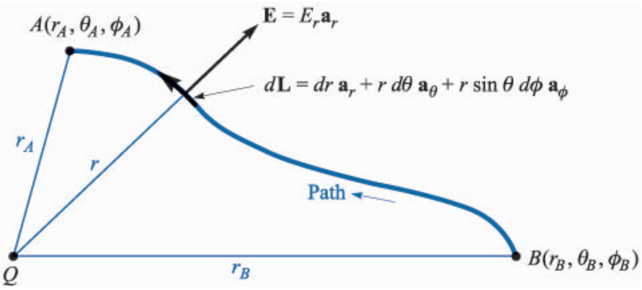


FIGURE 4.3
A general path between general points B and A in the field of a point charge Q at the origin. The potential difference V_{AB} is independent of the path selected.

each point from the charge and does not depend on the particular path used to carry our unit charge from one point to the other.

How might we conveniently define a zero reference for potential? The simplest possibility is to let $V = 0$ at infinity. If we let the point at $r = r_B$ recede to infinity the potential at r_A becomes

$$V_A = \frac{Q}{4\pi\epsilon_0 r_A}$$

or, since there is no reason to identify this point with the A subscript,

$$\boxed{V = \frac{Q}{4\pi\epsilon_0 r}} \tag{15}$$

This expression defines the potential at any point distant r from a point charge Q at the origin, the potential at infinite radius being taken as the zero reference. Returning to a physical interpretation, we may say that $Q/4\pi\epsilon_0 r$ joules of work must be done in carrying a 1-C charge from infinity to any point r meters from the charge Q.

A convenient method to express the potential without selecting a specific zero reference entails identifying r_A as r once again and letting $Q/4\pi\epsilon_0 r_B$ be a constant. Then

$$V = \frac{Q}{4\pi\epsilon_0 r} + C_1 \tag{16}$$

and C_1 may be selected so that $V = 0$ at any desired value of r. We could also select the zero reference indirectly by electing to let V be V_0 at $r = r_0$.

It should be noted that the *potential difference* between two points is not a function of C_1.

Equation (15) or (16) represents the potential field of a point charge. The potential is a scalar field and does not involve any unit vectors.

Let us now define an *equipotential surface* as a surface composed of all those points having the same value of potential. No work is involved in moving a unit charge around on an equipotential surface, for, by definition, there is no potential difference between any two points on this surface.

The equipotential surfaces in the potential field of a point charge are spheres centered at the point charge.

An inspection of the form of the potential field of a point charge shows that it is an inverse-distance field, whereas the electric field intensity was found to be an inverse-square-law relationship. A similar result occurs for the gravitational force field of a point mass (inverse-square law) and the gravitational potential field (inverse distance). The gravitational force exerted by the earth on an object one million miles from it is four times that exerted on the same object two million miles away. The kinetic energy given to a freely falling object starting from the end of the universe with zero velocity, however, is only twice as much at one million miles as it is at two million miles.

✔ **D4.5.** A 15-nC point charge is at the origin in free space. Calculate V_1 if point P_1 is located at $P_1(-2, 3, -1)$ and: (*a*) $V = 0$ at $(6, 5, 4)$; (*b*) $V = 0$ at infinity; (*c*) $V = 5\,\text{V}$ at $(2, 0, 4)$

Ans. 20.7 V; 36.0 V; 10.89 V

4.5 THE POTENTIAL FIELD OF A SYSTEM OF CHARGES: CONSERVATIVE PROPERTY

The potential at a point has been defined as the work done in bringing a unit positive charge from the zero reference to the point, and we have suspected that this work, and hence the potential, is independent of the path taken. If it were not, potential would not be a very useful concept.

Let us now prove our assertion. We shall do so by beginning with the potential field of the single point charge for which we showed, in the last section, the independence with regard to the path, noting that the field is linear with respect to charge so that superposition is applicable. It will then follow that the potential of a system of charges has a value at any point which is independent of the path taken in carrying the test charge to that point.

Thus the potential field of a single point charge, which we shall identify as Q_1 and locate at $\mathbf{r}_1$, involves only the distance $|\mathbf{r} - \mathbf{r}_1|$ from Q_1 to the point at $\mathbf{r}$ where we are establishing the value of the potential. For a zero reference at infinity, we have

$$V(\mathbf{r}) = \frac{Q_1}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}_1|}$$

The potential due to two charges, Q_1 at $\mathbf{r}_1$ and Q_2 at $\mathbf{r}_2$, is a function only of $|\mathbf{r} - \mathbf{r}_1|$ and $|\mathbf{r} - \mathbf{r}_2|$, the distances from Q_1 and Q_2 to the field point, respectively.

$$V(\mathbf{r}) = \frac{Q_1}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_1|} + \frac{Q_2}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_2|}$$

Continuing to add charges, we find that the potential due to n point charges is

$$V(\mathbf{r}) = \frac{Q_1}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_1|} + \frac{Q_2}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_2|} + \ldots + \frac{Q_n}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_n|}$$

or

$$V(\mathbf{r}) = \sum_{m=1}^{n} \frac{Q_m}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_m|} \tag{17}$$

If each point charge is now represented as a small element of a continuous volume charge distribution $\rho_v \Delta v$, then

$$V(\mathbf{r}) = \frac{\rho_v(\mathbf{r}_1)\Delta v_1}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_1|} + \frac{\rho_v(\mathbf{r}_2)\Delta v_2}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_2|} + \ldots + \frac{\rho_v(\mathbf{r}_n)\Delta v_n}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_n|}$$

As we allow the number of elements to become infinite, we obtain the integral expression

$$V(\mathbf{r}) = \int_{\text{vol}} \frac{\rho_v(\mathbf{r}')dv'}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|} \tag{18}$$

We have come quite a distance from the potential field of the single point charge, and it might be helpful to examine (18) and refresh ourselves as to the meaning of each term. The potential $V(\mathbf{r})$ is determined with respect to a zero reference potential at infinity and is an exact measure of the work done in bringing a unit charge from infinity to the field point at $\mathbf{r}$ where we are finding the potential. The volume charge density $\rho_y(\mathbf{r}')$ and differential volume element dv' combine to represent a differential amount of charge $\rho_v(\mathbf{r}')\,dv'$ located at $\mathbf{r}'$. The distance $|\mathbf{r}-\mathbf{r}'|$ is that distance from the source point to the field point. The integral is a multiple (volume) integral.

If the charge distribution takes the form of a line charge or a surface charge, the integration is along the line or over the surface:

$$V(\mathbf{r}) = \int \frac{\rho_L(\mathbf{r}')dL'}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|} \tag{19}$$

$$V(\mathbf{r}) = \int_S \frac{\rho_S(\mathbf{r}')dS'}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|} \tag{20}$$

The most general expression for potential is obtained by combining (17), (18), (19), and (20).

These integral expressions for potential in terms of the charge distribution should be compared with similar expressions for the electric field intensity, such as (18) in Sec. 2.3:

$$\mathbf{E}(\mathbf{r}) = \int_{\text{vol}} \frac{\rho_v(\mathbf{r}')dv'}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|^2} \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}$$

The potential again is inverse distance, and the electric field intensity, inverse-square law. The latter, of course, is also a vector field.

To illustrate the use of one of these potential integrals, let us find V on the z axis for a uniform line charge ρ_L in the form of a ring, $\rho = a$, in the $z = 0$ plane, as shown in Fig. 4.4. Working with (19), we have $dL' = a\,d\phi'$, $\mathbf{r} = z\mathbf{a}_z$, $\mathbf{r}' = a\mathbf{a}_\rho$, $|\mathbf{r} - \mathbf{r}'| = \sqrt{a^2 + z^2}$, and

$$V = \int_0^{2\pi} \frac{\rho_L a\, d\phi'}{4\pi\epsilon_0\sqrt{a^2 + z^2}} = \frac{\rho_L a}{2\pi\epsilon_0\sqrt{a^2 + z^2}}$$

For a zero reference at infinity, then:

1. The potential due to a single point charge is the work done in carrying a unit positive charge from infinity to the point at which we desire the potential, and the work is independent of the path chosen between those two points.
2. The potential field in the presence of a number of point charges is the sum of the individual potential fields arising from each charge.
3. The potential due to a number of point charges or any continuous charge distribution may therefore be found by carrying a unit charge from infinity to the point in question along any path we choose.

In other words, the expression for potential (zero reference at infinity),

$$V_A = -\int_\infty^A \mathbf{E} \cdot d\mathbf{L}$$

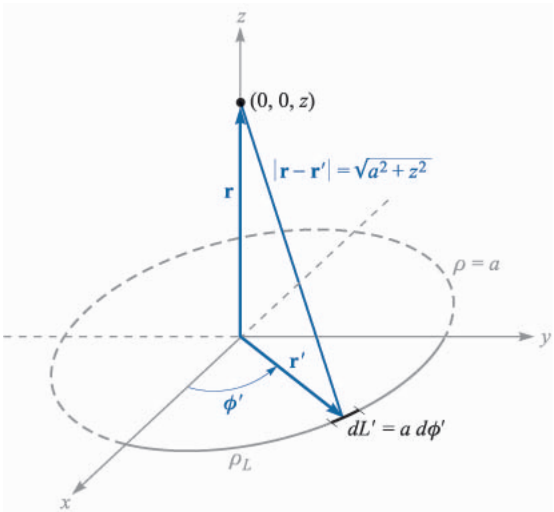


FIGURE 4.4
The potential field of a ring of uniform line charge density is easily obtained from $V = \int \rho_L(\mathbf{r}')\, dL' / (4\pi\epsilon_0|\mathbf{r} - \mathbf{r}'|)$.

or potential difference,

$$V_{AB} = V_A - V_B = -\int_B^A \mathbf{E} \cdot d\mathbf{L}$$

is not dependent on the path chosen for the line integral, regardless of the source of the **E** field.

This result is often stated concisely by recognizing that no work is done in carrying the unit charge around any *closed path*, or

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0 \tag{21}$$

A small circle is placed on the integral sign to indicate the closed nature of the path. This symbol also appeared in the formulation of Gauss's law, where a closed *surface* integral was used.

Equation (21) is true for *static* fields, but we shall see in Chap. 10 that Faraday demonstrated it was incomplete when time-varying magnetic fields were present. One of Maxwell's greatest contributions to electromagnetic theory was in showing that a time-varying electric field produces a magnetic field, and therefore we should expect to find later that (21) is not correct when either **E** or the magnetic field varies with time.

Restricting our attention to the static case where **E** does not change with time, consider the dc circuit shown in Fig. 4.5. Two points, A and B, are marked, and (21) states that no work is involved in carrying a unit charge from A through R_2 and R_3 to B and back to A through R_1, or that the sum of the potential differences around any closed path is zero.

Equation (21) is therefore just a more general form of Kirchhoff's circuital law for voltages, more general in that we can apply it to any region where an electric field exists and we are not restricted to a conventional circuit composed of wires, resistances, and batteries. Equation (21) must be amended before we can apply it to time-varying fields. We shall take care of this in Chap. 10, and in Chap. 13 we will then be able to establish the general form of Kirchhoff's voltage law for circuits in which currents and voltages vary with time.

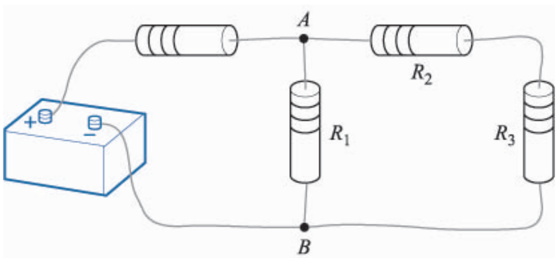


FIGURE 4.5
A simple dc-circuit problem which must be solved by applying $\oint \mathbf{E} \cdot d\mathbf{L} = 0$ in the form of Kirchhoff's voltage law.

Any field that satisfies an equation of the form of (21), (i.e., where the closed line integral of the field is zero) is said to be a *conservative field*. The name arises from the fact that no work is done (or that energy is *conserved*) around a closed path. The gravitational field is also conservative, for any energy expended in moving (raising) an object against the field is recovered exactly when the object is returned (lowered) to its original position. A nonconservative gravitational field could solve our energy problems forever.

Given a *nonconservative* field, it is of course possible that the line integral may be zero for certain closed paths. For example, consider the force field, $\mathbf{F} = \sin \pi\rho \, \mathbf{a}_\phi$. Around a circular path of radius $\rho = \rho_1$, we have $d\mathbf{L} = \rho \, d\phi \, \mathbf{a}_\phi$, and

$$\oint \mathbf{F} \cdot d\mathbf{L} = \int_0^{2\pi} \sin \pi\rho_1 \mathbf{a}_\phi \cdot \rho_1 d\phi \, \mathbf{a}_\phi = \int_0^{2\pi} \rho_1 \sin \pi\rho_1 \, d\phi$$

$$= 2\pi\rho_1 \sin \pi\rho_1$$

The integral is zero if $\rho_1 = 1, 2, 3, \ldots$, etc., but it is not zero for other values of ρ_1, or for most other closed paths, and the given field is not conservative. A conservative field must yield a zero value for the line integral around every possible closed path.

✔ **D4.6.** If we take the zero reference for potential at infinity, find the potential at (0, 0, 2) caused by this charge configuration in free space: (*a*) 12 nC/m on the line $\rho = 2.5$ m, $z = 0$; (*b*) point charge of 18 nC at $(1, 2, -1)$; (*c*) 12 nC/m on the line $y = 2.5$, $z = 0$.

Ans. 529 V; 43.2 V; 67.4 V

4.6 POTENTIAL GRADIENT

We now have two methods of determining potential, one directly from the electric field intensity by means of a line integral, and another from the basic charge distribution itself by a volume integral. Neither method is very helpful in determining the fields in most practical problems, however, for as we shall see later, neither the electric field intensity nor the charge distribution is very often known. Preliminary information is much more apt to consist of a description of two equipotential surfaces, such as the statement that we have two parallel conductors of circular cross section at potentials of 100 and -100 V. Perhaps we wish to find the capacitance between the conductors, or the charge and current distribution on the conductors from which losses may be calculated.

These quantities may be easily obtained from the potential field, and our immediate goal will be a simple method of finding the electric field intensity from the potential.

We already have the general line-integral relationship between these quantities,

$$V = -\int \mathbf{E} \cdot d\mathbf{L} \tag{22}$$

but this is much easier to use in the reverse direction: given **E**, find V.

However, (22) may be applied to a very short element of length $\Delta\mathbf{L}$ along which **E** is essentially constant, leading to an incremental potential difference ΔV,

$$\Delta V \doteq -\mathbf{E} \cdot \Delta\mathbf{L} \tag{23}$$

Let us see first if we can determine any new information about the relation of V to **E** from this equation. Consider a general region of space, as shown in Fig. 4.6, in which **E** and V both change as we move from point to point. Equation (23) tells us to choose an incremental vector element of length $\Delta\mathbf{L} = \Delta L\, \mathbf{a}_L$ and multiply its magnitude by the component of **E** in the direction of $\mathbf{a}_L$ (one interpretation of the dot product) to obtain the small potential difference between the final and initial points of $V\mathbf{L}$.

If we designate the angle between $\Delta\mathbf{L}$ and **E** as θ, then

$$\Delta V \doteq -E\Delta L \cos\theta$$

We now wish to pass to the limit and consider the derivative dV/dL. To do this, we need to show that V may be interpreted as a *function* $V(x, y, z)$. So far, V is merely the result of the line integral (22). If we assume a specified starting point or zero reference and then let our end point be (x, y, z), we know that the result of the integration is a unique function of the end point (x, y, z) because **E** is a conservative field. Therefore V is a single-valued function $V(x, y, z)$. We may then pass to the limit and obtain

$$\frac{dV}{dL} = -E\cos\theta$$

In which direction should $\Delta\mathbf{L}$ be placed to obtain a maximum value of ΔV? Remember that **E** is a definite value at the point at which we are working and is independent of the direction of $\Delta\mathbf{L}$. The magnitude ΔL is also constant, and our

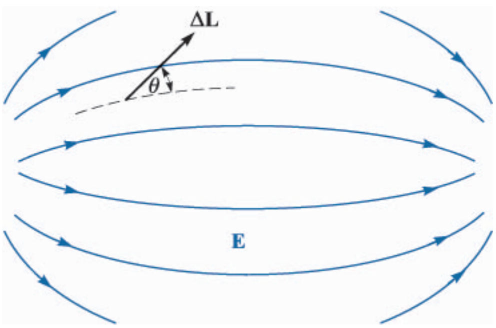


FIGURE 4.6
A vector incremental element of length $\Delta\mathbf{L}$ is shown making an angle of θ with an **E** field, indicated by its streamlines. The sources of the field are not shown.

variable is $\mathbf{a}_L$, the unit vector showing the direction of $\Delta\mathbf{L}$. It is obvious that the maximum positive increment of potential, $\Delta V_{\max}$, will occur when $\cos\theta$ is -1, or $\Delta\mathbf{L}$ points in the direction *opposite* to $\mathbf{E}$. For this condition,

$$\left.\frac{dV}{dL}\right|_{\max} = E$$

This little exercise shows us two characteristics of the relationship between $\mathbf{E}$ and V at any point:

1. The magnitude of the electric field intensity is given by the maximum value of the rate of change of potential with distance.
2. This maximum value is obtained when the direction of the distance increment is opposite to $\mathbf{E}$ or, in other words, the direction of $\mathbf{E}$ is *opposite* to the direction in which the potential is *increasing* the most rapidly.

Let us now illustrate these relationships in terms of potential. Fig. 4.7 is intended to show the information we have been given about some potential field. It does this by showing the equipotential surfaces (shown as lines in the two-dimensional sketch). We desire information about the electric field intensity at point P. Starting at P, we lay off a small incremental distance $\Delta\mathbf{L}$ in various directions, hunting for that direction in which the potential is changing (increasing) the most rapidly. From the sketch, this direction appears to be left and slightly upward. From our second characteristic above, the electric field intensity is therefore oppositely directed, or to the right and slightly downward at P. Its magnitude is given by dividing the small increase in potential by the small element of length.

It seems likely that the direction in which the potential is increasing the most rapidly is perpendicular to the equipotentials (in the direction of *increasing* potential), and this is correct, for if $\Delta\mathbf{L}$ is directed along an equipotential, $\Delta V = 0$ by our definition of an equipotential surface. But then

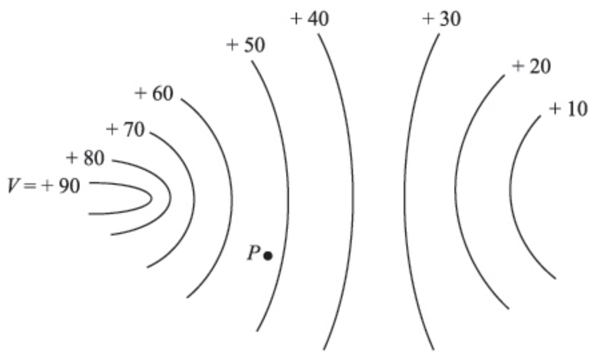


FIGURE 4.7
A potential field is shown by its equipotential surfaces. At any point the **E** field is normal to the equipotential surface passing through that point and is directed toward the more negative surfaces.

$$\Delta V = -\mathbf{E} \cdot \Delta \mathbf{L} = 0$$

and since neither $\mathbf{E}$ nor $\Delta\mathbf{L}$ is zero, $\mathbf{E}$ must be perpendicular to this $\Delta\mathbf{L}$ or perpendicular to the equipotentials.

Since the potential field information is more likely to be determined first, let us describe the direction of $\Delta\mathbf{L}$ which leads to a maximum increase in potential mathematically in terms of the potential field rather than the electric field intensity. We do this by letting $\mathbf{a}_N$ be a unit vector normal to the equipotential surface and directed toward the higher potentials. The electric field intensity is then expressed in terms of the potential,

$$\mathbf{E} = -\left.\frac{dV}{dL}\right|_{\max} \mathbf{a}_N \tag{24}$$

which shows that the magnitude of $\mathbf{E}$ is given by the maximum space rate of change of V and the direction of $\mathbf{E}$ is *normal* to the equipotential surface (in the direction of *decreasing* potential).

Since $dV/dL|_{\max}$ occurs when $\Delta\mathbf{L}$ is in the direction of $\mathbf{a}_N$, we may remind ourselves of this fact by letting

$$\left.\frac{dV}{dL}\right|_{\max} = \frac{dV}{dN}$$

and

$$\mathbf{E} = -\frac{dV}{dN}\mathbf{a}_N \tag{25}$$

Equation (24) or (25) serves to provide a physical interpretation of the process of finding the electric field intensity from the potential. Both are descriptive of a general procedure, and we do not intend to use them directly to obtain quantitative information. This procedure leading from V to $\mathbf{E}$ is not unique to this pair of quantities, however, but has appeared as the relationship between a scalar and a vector field in hydraulics, thermodynamics, and magnetics, and indeed in almost every field to which vector analysis has been applied.

The operation on V by which $-\mathbf{E}$ is obtained is known as the *gradient*, and the gradient of a scalar field T is defined as

$$\boxed{\text{Gradient of } T = \text{grad } T = \frac{dT}{dN}\mathbf{a}_N} \tag{26}$$

where $\mathbf{a}_N$ is a unit vector normal to the equipotential surfaces, and that normal is chosen which points in the direction of increasing values of T.

Using this new term, we now may write the relationship between V and $\mathbf{E}$ as

$$\boxed{\mathbf{E} = -\text{grad } V} \tag{27}$$

Since we have shown that V is a unique function of x, y, and z, we may take its total differential

$$dV = \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz$$

But we also have

$$dV = -\mathbf{E} \cdot d\mathbf{L} = -E_x\, dx - E_y\, dy - E_z\, dz$$

Since both expressions are true for any dx, dy, and dz, then

$$E_x = -\frac{\partial V}{\partial x}$$
$$E_y = -\frac{\partial V}{\partial y}$$
$$E_z = -\frac{\partial V}{\partial z}$$

These results may be combined vectorially to yield

$$\boxed{\mathbf{E} = -\left(\frac{\partial V}{\partial x}\mathbf{a}_x + \frac{\partial V}{\partial y}\mathbf{a}_y + \frac{\partial V}{\partial z}\mathbf{a}_z\right)} \qquad (28)$$

and comparison of (27) and (28) provides us with an expression which may be used to evaluate the gradient in cartesian coordinates,

$$\boxed{\text{grad } V = \frac{\partial V}{\partial x}\mathbf{a}_x + \frac{\partial V}{\partial y}\mathbf{a}_y + \frac{\partial V}{\partial z}\mathbf{a}_z} \qquad (29)$$

The gradient of a scalar is a vector, and old quizzes show that the unit vectors which are often incorrectly added to the divergence expression appear to be those which were incorrectly removed from the gradient. Once the physical interpretation of the gradient, expressed by (26), is grasped as showing the maximum space rate of change of a scalar quantity and *the direction in which this maximum occurs*, the vector nature of the gradient should be self-evident.

The vector operator

$$\nabla = \frac{\partial}{\partial x}\mathbf{a}_x + \frac{\partial}{\partial y}\mathbf{a}_y + \frac{\partial}{\partial z}\mathbf{a}_z$$

may be used formally as an operator on a scalar, T, ∇T, producing

$$\nabla T = \frac{\partial T}{\partial x}\mathbf{a}_x + \frac{\partial T}{\partial y}\mathbf{a}_y + \frac{\partial T}{\partial z}\mathbf{a}_z$$

from which we see that

$$\boxed{\nabla T = \text{grad } T}$$

This allows us to use a very compact expression to relate **E** and V,

$$\boxed{\mathbf{E} = -\nabla V} \tag{30}$$

The gradient may be expressed in terms of partial derivatives in other coordinate systems through application of its definition (26). These expressions are derived in Appendix A and repeated below for convenience when dealing with problems having cylindrical or spherical symmetry. They also appear inside the back cover.

$$\boxed{\nabla V = \frac{\partial V}{\partial x}\mathbf{a}_x + \frac{\partial V}{\partial y}\mathbf{a}_y + \frac{\partial V}{\partial z}\mathbf{a}_z \qquad \text{(cartesian)}} \tag{31}$$

$$\boxed{\nabla V = \frac{\partial V}{\partial \rho}\mathbf{a}_\rho + \frac{1}{\rho}\frac{\partial V}{\partial \phi}\mathbf{a}_\phi + \frac{\partial V}{\partial z}\mathbf{a}_z \qquad \text{(cylindrical)}} \tag{32}$$

$$\boxed{\nabla V = \frac{\partial V}{\partial r}\mathbf{a}_r + \frac{1}{r}\frac{\partial V}{\partial \theta}\mathbf{a}_\theta + \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi}\mathbf{a}_\phi \qquad \text{(spherical)}} \tag{33}$$

Note that the denominator of each term has the form of one of the components of $d\mathbf{L}$ in that coordinate system, except that partial differentials replace ordinary differentials; for example, $r\sin\theta\, d\phi$ becomes $r\sin\theta\, \partial\phi$.

Let us now hasten to illustrate the gradient concept with an example.

Example 4.3

Given the potential field, $V = 2x^2y - 5z$, and a point $P(-4, 3, 6)$, we wish to find several numerical values at point P: the potential V, the electric field intensity **E**, the direction of **E**, the electric flux density **D**, and the volume charge density ρ_v.

Solution. The potential at $P(-4, 5, 6)$ is

$$V_P = 2(-4)^2(3) - 5(6) = 66 \quad \text{V}$$

Next, we may use the gradient operation to obtain the electric field intensity,

$$\mathbf{E} = -\nabla V = -4xy\mathbf{a}_x - 2x^2\mathbf{a}_y + 5\mathbf{a}_z \quad \text{V/m}$$

The value of $\mathbf{E}$ at point P is

$$\mathbf{E}_P = 48\mathbf{a}_x - 32\mathbf{a}_y + 5\mathbf{a}_z \quad \text{V/m}$$

and

$$|\mathbf{E}_P| = \sqrt{48^2 + (-32)^2 + 5^2} = 57.9 \quad \text{V/m}$$

The direction of $\mathbf{E}$ at P is given by the unit vector

$$\mathbf{a}_{E,P} = (48\mathbf{a}_x - 32\mathbf{a}_y + 5\mathbf{a}_z)/57.9$$
$$= 0.829\mathbf{a}_x - 0.553\mathbf{a}_y + 0.086\mathbf{a}_z$$

If we assume these fields exist in free space, then

$$\mathbf{D} = \epsilon_0\mathbf{E} = -35.4xy\mathbf{a}_x - 17.71x^2\mathbf{a}_y + 44.3\mathbf{a}_z \quad \text{pC/m}^3$$

Finally, we may use the divergence relationship to find the volume charge density that is the source of the given potential field,

$$\rho_v = \nabla \cdot \mathbf{D} = -35.4y \quad \text{pC/m}^3$$

At P, $\rho_v = -106.2$ pC/m^3.

✔ **D4.7.** A portion of a two-dimensional ($E_z = 0$) potential field is shown in Fig. 4.8. The grid lines are 1 mm apart in the actual field. Determine approximate values for $\mathbf{E}$ in cartesian coordinates at: (*a*) *a*; (*b*) *b*; (*c*) *c*.

Ans. $-1075\mathbf{a}_y$ V/m; $-600\mathbf{a}_x - 700\mathbf{a}_y$ V/m; $-500\mathbf{a}_x - 650\mathbf{a}_y$ V/m

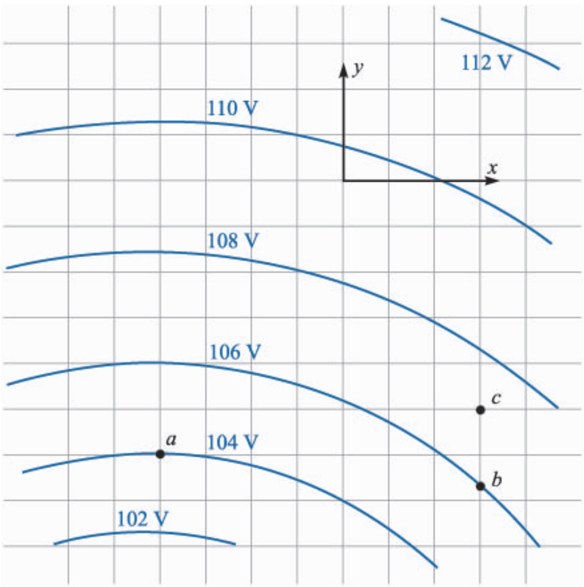


FIGURE 4.8
See Prob. D4.7.

✔ **D4.8.** Given the potential field in cylindrical coordinates, $V = \frac{100}{z^2+1}\rho\cos\phi$ V, and point P at $\rho = 3$ m, $\phi = 60^\circ$, $z = 2$ m, find values at P for: (*a*) V; (*b*) $\mathbf{E}$; (*c*) E; (*d*) dV/dN; (*e*) $\mathbf{a}_N$; (*f*) ρ_v in free space.

Ans. 30.0 V; $-10.00\mathbf{a}_\rho + 8.66\mathbf{a}_\phi + 24.0\mathbf{a}_z$ V/m; 27.4 V/m; $-0.365\mathbf{a}_\rho - 0.316\mathbf{a}_\phi - 0.876\mathbf{a}_z$; 27.4 V/m; -234 pC/m^3.

4.7 THE DIPOLE

The dipole fields which we shall develop in this section are quite important because they form the basis for the behavior of dielectric materials in electric fields, as discussed in part of the following chapter, as well as justifying the use of images, as described in Sec. 5.5 of the next chapter. Moreover, this development will serve to illustrate the importance of the potential concept presented in this chapter.

An *electric dipole*, or simply a *dipole*, is the name given to two point charges of equal magnitude and opposite sign, separated by a distance which is small compared to the distance to the point P at which we want to know the electric and potential fields. The dipole is shown in Fig. 4.9*a*. The distant point P is described by the spherical coordinates r, θ, and $\phi - 90^\circ$, in view of the azimuthal symmetry. The positive and negative point charges have separation d and cartesian coordinates $(0, 0, \frac{1}{2}d)$ and $(0, 0, -\frac{1}{2}d)$, respectively.

So much for the geometry. What would we do next? Should we find the total electric field intensity by adding the known fields of each point charge? Would it be easier to find the total potential field first? In either case, having found one, we shall find the other from it before calling the problem solved.

If we choose to find $\mathbf{E}$ first, we shall have two components to keep track of in spherical coordinates (symmetry shows E_ϕ is zero), and then the only way to find V from $\mathbf{E}$ is by use of the line integral. This last step includes establishing a suitable zero reference for potential, since the line integral gives us only the potential difference between the two points at the ends of the integral path.

$$V = \frac{Q}{4\pi\epsilon_0}\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \frac{Q}{4\pi\epsilon_0}\frac{R_2 - R_1}{R_1 R_2}$$

Note that the plane $z = 0$, midway between the two point charges, is the locus of points for which $R_1 = R_2$, and is therefore at zero potential, as are all points at infinity.

For a distant point, $R_1 \doteq R_2$, and the $R_1 R_2$ product in the denominator may be replaced by r^2. The approximation may not be made in the numerator, however, without obtaining the trivial answer that the potential field approaches zero as we go very far away from the dipole. Coming back a little closer to the dipole, we see from Fig. 4.9*b* that $R_2 - R_1$ may be approximated very easily if R_1 and R_2 are assumed to be parallel,

$$R_2 - R_1 \doteq d\cos\theta$$

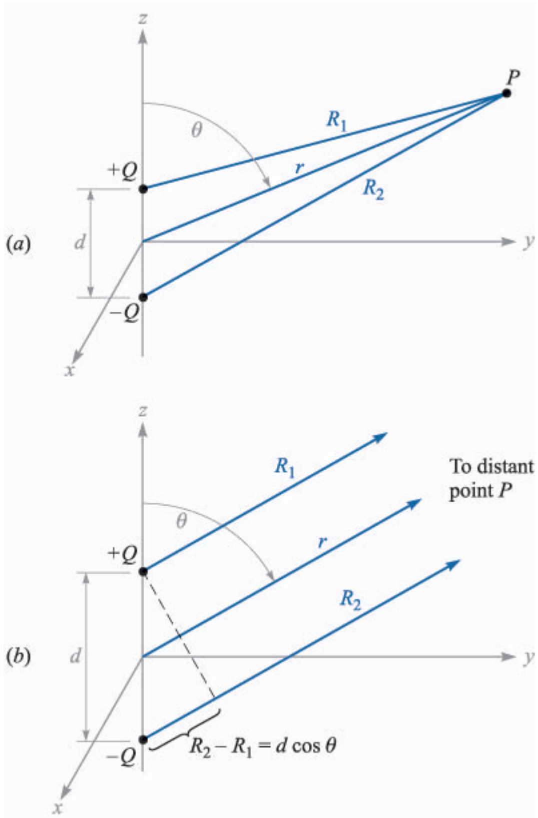


FIGURE 4.9
(*a*) The geometry of the problem of an electric dipole. The dipole moment $\mathbf{p} = Q\mathbf{d}$ is in the $\mathbf{a}_z$ direction. (*b*) For a distant point P, R_1 is essentially parallel to R_2, and we find that $R_2 - R_1 = d\cos\theta$.

The final result is then

$$\boxed{V = \frac{Qd\cos\theta}{4\pi\epsilon_0 r^2}} \tag{34}$$

Again we note that the plane $z = 0$ ($\theta = 90^\circ$) is at zero potential.

Using the gradient relationship in spherical coordinates,

$$\mathbf{E} = -\nabla V = -\left(\frac{\partial V}{\partial r}\mathbf{a}_r + \frac{1}{r}\frac{\partial V}{\partial \theta}\mathbf{a}_\theta + \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi}\mathbf{a}_\phi\right)$$

we obtain

$$\mathbf{E} = -\left(-\frac{Qd\cos\theta}{4\pi\epsilon_0 r^3}\mathbf{a}_r - \frac{Qd\sin\theta}{4\pi\epsilon_0 r^3}\mathbf{a}_\theta\right) \tag{35}$$

or

$$\mathbf{E} = \frac{Qd}{4\pi\epsilon_0 r^3}(2\cos\theta\,\mathbf{a}_r + \sin\theta\,\mathbf{a}_\theta) \tag{36}$$

These are the desired distant fields of the dipole, obtained with a very small amount of work. Any student who has several hours to spend may try to work the problem in the reverse direction—the authors consider the process too long and detailed to include, even for effect.

To obtain a plot of the potential field, we may choose a dipole such that $Qd/(4\pi\epsilon_0) = 1$, and then $\cos\theta = Vr^2$. The colored lines in Fig. 4.10 indicate equipotentials for which $V = 0, +0.2, +0.4, +0.6, +0.8$, and $+1$, as indicated. The dipole axis is vertical, with the positive charge on the top. The streamlines for the electric field are obtained by applying the methods of Sec. 2.6 in spherical coordinates,

$$\frac{E_\theta}{E_r} = \frac{r\,d\theta}{dr} = \frac{\sin\theta}{2\cos\theta}$$

or

$$\frac{dr}{r} = 2\cot\theta\,d\theta$$

from which we obtain

$$r = C_1 \sin^2\theta$$

The black streamlines shown in Fig. 4.10 are for $C_1 = 1, 1.5, 2$, and 2.5.

The potential field of the dipole, Eq. (34), may be simplified by making use of the dipole moment. Let us first identify the vector length directed from $-Q$ to $+Q$ as $\mathbf{d}$ and then define the *dipole moment* as $Q\mathbf{d}$ and assign it the symbol $\mathbf{p}$. Thus

$$\mathbf{p} = Q\mathbf{d} \tag{37}$$

The units of $\mathbf{p}$ are $\mathrm{C \cdot m}$.

Since $\mathbf{d} \cdot \mathbf{a}_r = d\cos\theta$, we then have

$$V = \frac{\mathbf{p} \cdot \mathbf{a}_r}{4\pi\epsilon_0 r^2} \tag{38}$$

This result may be generalized as

$$V = \frac{1}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|^2}\,\mathbf{p} \cdot \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \tag{39}$$

where $\mathbf{r}$ locates the field point P, and $\mathbf{r}'$ determines the dipole center. Equation (39) is independent of any coordinate system.

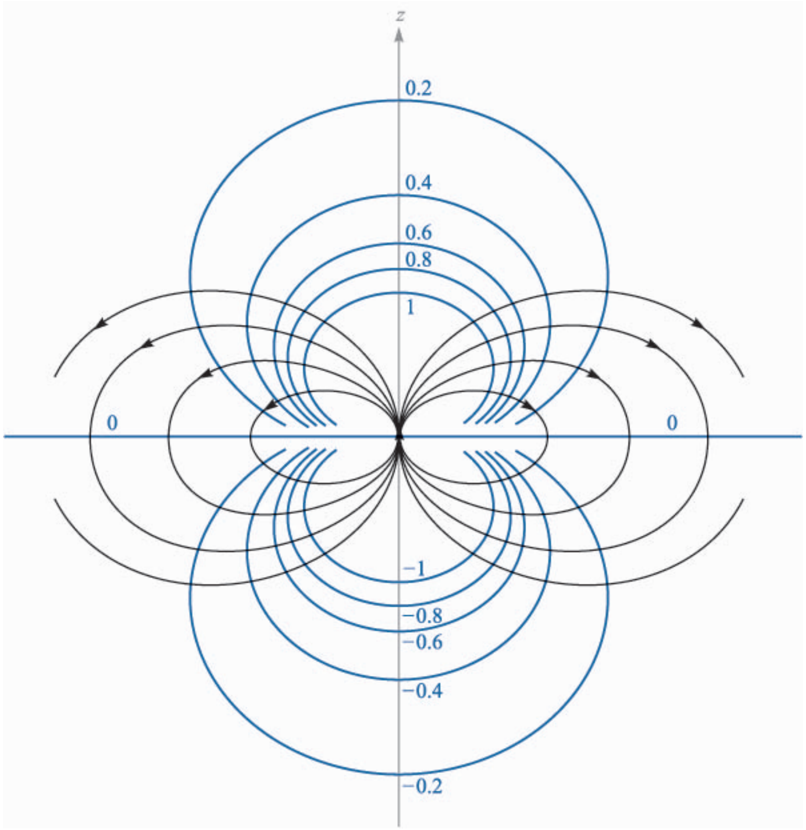


FIGURE 4.10
The electrostatic field of a point dipole with its moment in the $\mathbf{a}_z$ direction. Six equipotential surfaces are labeled with relative values of V.

The dipole moment **p** will appear again when we discuss dielectric materials. Since it is equal to the product of the charge and the separation, neither the dipole moment nor the potential will change as Q increases and **d** decreases, provided the product remains constant. The limiting case of a *point dipole* is achieved when we let **d** approach zero and Q approach infinity such that the product **p** is finite.

Turning our attention to the resultant fields, it is interesting to note that the potential field is now proportional to the inverse *square* of the distance, and the electric field intensity is proportional to the inverse *cube* of the distance from the dipole. Each field falls off faster than the corresponding field for the point charge, but this is no more than we should expect because the opposite charges appear to be closer together at greater distances and to act more like a single point charge of 0 C.

Symmetrical arrangements of larger numbers of point charges produce fields proportional to the inverse of higher and higher powers of r. These charge distributions are called *multipoles*, and they are used in infinite series to approximate more unwieldy charge configurations.

✔ **D4.9.** An electric dipole located at the origin in free space has a moment $\mathbf{p} = 3\mathbf{a}_x - 2\mathbf{a}_y + \mathbf{a}_z$ nC · m. (*a*) Find V at $P_A(2, 3, 4)$. (*b*) Find V at $r = 2.5$, $\theta = 30^\circ$, $\phi = 40^\circ$.

Ans. 0.230 V; 1.973 V

✔ **D4.10.** A dipole of moment $\mathbf{p} = 6\mathbf{a}_z$ nC · m is located at the origin in free space. (*a*) Find V at $P(r = 4, \theta = 20^\circ, \phi = 0^\circ)$. (*b*) Find $\mathbf{E}$ at P.

Ans. 3.17 V; $1.584\mathbf{a}_r + 0.288\mathbf{a}_\theta$ V/m

4.8 ENERGY DENSITY IN THE ELECTROSTATIC FIELD

We have introduced the potential concept by considering the work done, or energy expended, in moving a point charge around in an electric field, and now we must tie up the loose ends of that discussion by tracing the energy flow one step further.

Bringing a positive charge from infinity into the field of another positive charge requires work, the work being done by the external source moving the charge. Let us imagine that the external source carries the charge up to a point near the fixed charge and then holds it there. Energy must be conserved, and the energy expended in bringing this charge into position now represents potential energy, for if the external source released its hold on the charge, it would accelerate away from the fixed charge, acquiring kinetic energy of its own and the capability of doing work.

In order to find the potential energy present in a system of charges, we must find the work done by an external source in positioning the charges.

We may start by visualizing an empty universe. Bringing a charge Q_1 from infinity to any position requires no work, for there is no field present.[2] The positioning of Q_2 at a point in the field of Q_1 requires an amount of work given by the product of the charge Q_2 and the potential at that point due to Q_1. We represent this potential as $V_{2,1}$, where the first subscript indicates the location and the second subscript the source. That is, $V_{2,1}$ is the potential at the location of Q_2 due to Q_1. Then

$$\text{Work to position } Q_2 = Q_2 V_{2,1}$$

[2] However, somebody in the workshop at infinity had to do an infinite amount of work to create the point charge in the first place! How much energy is required to bring two half-charges into coincidence to make a unit charge?

Similarly, we may express the work required to position each additional charge in the field of all those already present:

$$\text{Work to position } Q_3 = Q_3 V_{3,1} + Q_3 V_{3,2}$$
$$\text{Work to position } Q_4 = Q_4 V_{4,1} + Q_4 V_{4,2} + Q_4 V_{4,3}$$

and so forth. The total work is obtained by adding each contribution:

$$\begin{aligned}\text{Total positioning work } &= \text{ potential energy of field} \\ &= W_E = Q_2 V_{2,1} + Q_3 V_{3,1} + Q_3 V_{3,2} + Q_4 V_{4,1} \\ &\quad + Q_4 V_{4,2} + Q_4 V_{4,3} + \ldots \end{aligned} \tag{40}$$

Noting the form of a representative term in the above equation,

$$Q_3 V_{3,1} = Q_3 \frac{Q_1}{4\pi\epsilon_0 R_{13}} = Q_1 \frac{Q_3}{4\pi\epsilon_0 R_{31}}$$

where R_{13} and R_{31} each represent the scalar distance between Q_1 and Q_3, we see that it might equally well have been written as $Q_1 V_{1,3}$. If each term of the total energy expression is replaced by its equal, we have

$$W_E = Q_1 V_{1,2} + Q_1 V_{1,3} + Q_2 V_{2,3} + Q_1 V_{1,4} + Q_2 V_{2,4} + Q_3 V_{3,4} + \ldots \tag{41}$$

Adding the two energy expressions (40) and (41) gives us a chance to simplify the result a little:

$$\begin{aligned} 2W_E = Q_1(&V_{1,2} + V_{1,3} + V_{1,4} + \ldots) \\ &+ Q_2(V_{2,1} + V_{2,3} + V_{2,4} + \ldots) \\ &+ Q_3(V_{3,1} + V_{3,2} + V_{3,4} + \ldots) \\ &+ \ldots \end{aligned}$$

Each sum of potentials in parentheses is the combined potential due to all the charges except for the charge at the point where this combined potential is being found. In other words,

$$V_{1,2} + V_{1,3} + V_{1,4} + \ldots = V_1$$

the potential at the location of Q_1 due to the presence of $Q_2, Q_3, \ldots$. We therefore have

$$W_E = \tfrac{1}{2}(Q_1 V_1 + Q_2 V_2 + Q_3 V_3 + \ldots) = \tfrac{1}{2} \sum_{m=1}^{m=N} Q_m V_m \tag{42}$$

In order to obtain an expression for the energy stored in a region of continuous charge distribution, each charge is replaced by $\rho_v dv$, and the summation becomes an integral,

$$W_E = \tfrac{1}{2} \int_{\text{vol}} \rho_v V \, dv \tag{43}$$

Equations (42) and (43) allow us to find the total potential energy present in a system of point charges or distributed volume charge density. Similar expressions may be easily written in terms of line or surface charge density. Usually we prefer to use (43) and let it represent all the various types of charge which may have to be considered. This may always be done by considering point charges, line charge density, or surface charge density as continuous distributions of volume charge density over very small regions. We shall illustrate such a procedure with an example shortly.

Before we undertake any interpretation of this result, we should consider a few lines of more difficult vector analysis and obtain an expression equivalent to (43) but written in terms of **E** and **D**.

We begin by making the expression a little bit longer. Using Maxwell's first equation, replace ρ_v by its equal $\nabla \cdot \mathbf{D}$ and make use of a vector identity which is true for any scalar function V and any vector function **D**,

$$\nabla \cdot (V\mathbf{D}) \equiv V(\nabla \cdot \mathbf{D}) + \mathbf{D} \cdot (\nabla V) \tag{44}$$

This may be proved readily by expansion in cartesian coordinates. We then have, successively,

$$W_E = \frac{1}{2}\int_{\text{vol}} \rho_v V dv = \frac{1}{2}\int_{\text{vol}} (\nabla \cdot \mathbf{D})V\, dv$$

$$= \frac{1}{2}\int_{\text{vol}} [\nabla \cdot (V\mathbf{D}) - \mathbf{D} \cdot (\nabla V)]\, dv$$

Using the divergence theorem from the last chapter, the first volume integral of the last equation is changed into a closed surface integral, where the closed surface surrounds the volume considered. This volume, first appearing in (43), must contain *every* charge, and there can then be no charges outside of the volume. We may therefore consider the volume as *infinite* in extent if we wish. We have

$$W_E = \frac{1}{2}\oint_S (V\mathbf{D}) \cdot d\mathbf{S} - \frac{1}{2}\int_{\text{vol}} \mathbf{D} \cdot (\nabla V)\, dv$$

The surface integral is equal to zero, for over this closed surface surrounding the universe we see that V is approaching zero at least as rapidly as $1/r$ (the charges look like a point charge from there), **D** is approaching zero at least as rapidly as $1/r^2$, while the differential element of surface, looking more and more like a portion of a sphere, is increasing only as r^2. The integrand therefore approaches zero at least as rapidly as $1/r$. In the limit the integrand and the integral are zero. Substituting $\mathbf{E} = -\nabla V$ in the remaining volume integral, we have our answer,

$$\boxed{W_E = \frac{1}{2}\int_{\text{vol}} \mathbf{D} \cdot \mathbf{E}\, dv = \frac{1}{2}\int_{\text{vol}} \epsilon_0 E^2\, dv} \tag{45}$$

Let us now use this last expression to calculate the energy stored in the electrostatic field of a section of a coaxial cable or capacitor of length L. We found in Sec. 3.3 of the previous chapter that

$$D_\rho = \frac{a\rho_S}{\rho}$$

Hence,

$$\mathbf{E} = \frac{a\rho_S}{\epsilon_0 \rho}\mathbf{a}_\rho$$

where ρ_S is the surface charge density on the inner conductor, whose radius is a. Thus,

$$W_E = \tfrac{1}{2}\int_0^L \int_0^{2\pi} \int_a^b \epsilon_0 \frac{a^2\rho_S^2}{\epsilon_0^2\rho^2}\rho\, d\rho\, d\phi\, dz = \frac{\pi L a^2 \rho_S^2}{\epsilon_0}\ln\frac{b}{a}$$

This same result may be obtained from (43). We choose the outer conductor as our zero-potential reference, and the potential of the inner cylinder is then

$$V_a = -\int_b^a E_\rho\, d\rho = -\int_b^a \frac{a\rho_S}{\epsilon_0\rho}d\rho = \frac{a\rho_S}{\epsilon_0}\ln\frac{b}{a}$$

The surface charge density ρ_S at $\rho = a$ can be interpreted as a volume charge density $\rho_v = \rho_S/t$, extending from $\rho = a - \frac{1}{2}t$ to $\rho = a + \frac{1}{2}t$, where $t \ll a$. The integrand in (43) is therefore zero everywhere between the cylinders (where the volume charge density is zero), as well as at the outer cylinder (where the potential is zero). The integration is therefore performed only within the thin cylindrical shell at $\rho = a$,

$$W_E = \tfrac{1}{2}\int_{\text{vol}} \rho_v V\, dV = \tfrac{1}{2}\int_0^L \int_0^{2\pi} \int_{a-t/2}^{a+t/2} \frac{\rho_S}{t} a \frac{\rho_S}{\epsilon_0}\ln\frac{b}{a}\rho\, d\rho\, d\phi\, dz$$

from which

$$W_E = \frac{a^2\rho_S^2 \ln(b/a)}{\epsilon_0}\pi L$$

once again.

This expression takes on a more familiar form if we recognize the total charge on the inner conductor as $Q = 2\pi a L \rho_S$. Combining this with the potential difference between the cylinders, V_a, we see that

$$W_E = \tfrac{1}{2}QV_a$$

which should be familiar as the energy stored in a capacitor.

The question of where the energy is stored in an electric field has not yet been answered. Potential energy can never be pinned down precisely in terms of physical location. Someone lifts a pencil, and the pencil acquires potential energy. Is the energy stored in the molecules of the pencil, in the gravitational

field between the pencil and the earth, or in some obscure place? Is the energy in a capacitor stored in the charges themselves, in the field, or where? No one can offer any proof for his or her own private opinion, and the matter of deciding may be left to the philosophers.

Electromagnetic field theory makes it easy to believe that the energy of an electric field or a charge distribution is stored in the field itself, for if we take (45), an exact and rigorously correct expression,

$$W_E = \frac{1}{2}\int_{\text{vol}} \mathbf{D}\cdot\mathbf{E}\, dv$$

and write it on a differential basis,

$$dW_E = \tfrac{1}{2}\mathbf{D}\cdot\mathbf{E}\, dv$$

or

$$\frac{dW_E}{dv} = \tfrac{1}{2}\mathbf{D}\cdot\mathbf{E} \tag{46}$$

we obtain a quantity $\frac{1}{2}\mathbf{D}\cdot\mathbf{E}$, which has the dimensions of an energy density, or joules per cubic meter. We know that if we integrate this energy density over the entire field-containing volume, the result is truly the total energy present, but we have no more justification for saying that the energy stored in each differential volume element dv is $\mathbf{D}\cdot\mathbf{E}\, dv$ than we have for looking at (43) and saying that the stored energy is $\frac{1}{2}\rho_v V dv$. The interpretation afforded by (46), however, is a convenient one, and we shall use it until proved wrong.

✔ **D4.11.** Find the energy stored in free space for the region $2\,\text{mm} < r < 3\,\text{mm}$, $0 < \theta < 90^\circ$, $0 < \phi < 90^\circ$, given the potential field $V =$: (a) $\dfrac{200}{r}$ V; (b) $\dfrac{300\cos\theta}{r^2}$ V.

Ans. 1.391 pJ; 36.7 J

SUGGESTED REFERENCES

1. Attwood, S. S.: "Electric and Magnetic Fields," 3d ed., John Wiley & Sons, Inc., New York, 1949. There are a large number of well-drawn field maps of various charge distributions, including the dipole field. Vector analysis is not used.
2. Skilling, H. H.: (see Suggested References for Chap. 3). Gradient is described on pp. 19–21.
3. Thomas, G. B., Jr., and R. L. Finney: (see Suggested References for Chap. 1). The directional derivative and the gradient are presented on pp. 823–830.

PROBLEMS

4.1 The value of $\mathbf{E}$ at $P(\rho = 2,\ \phi = 40^\circ,\ z = 3)$ is given as $\mathbf{E} = 100\mathbf{a}_\rho - 200\mathbf{a}_\phi + 300\mathbf{a}_z$. V/m. Determine the incremental work required to move a 20-μC charge a distance of 6 μm in the direction of: (*a*) $\mathbf{a}_\rho$; (*b*) $\mathbf{a}_\phi$; (*c*) $\mathbf{a}_z$; (*d*) $\mathbf{E}$; (*e*) $\mathbf{G} = 2\mathbf{a}_x - 3\mathbf{a}_y + 4\mathbf{a}_z$.

4.2 Let $\mathbf{E} = 400\mathbf{a}_x - 300\mathbf{a}_y + 500\mathbf{a}_z$ V/m in the neighborhood of point $P(6, 2, -3)$. Find the incremental work done in moving a 4-C charge a distance of 1 mm in the direction specified by: (*a*) $\mathbf{a}_x + \mathbf{a}_y + \mathbf{a}_z$; (*b*) $-2\mathbf{a}_x + 3\mathbf{a}_y - \mathbf{a}_z$.

4.3 If $\mathbf{E} = 120\mathbf{a}_\rho$ V/m, find the incremental amount of work done in moving a 50-μC charge a distance of 2 mm from: (*a*) $P(1, 2, 3)$ toward $Q(2, 1, 4)$; (*b*) $Q(2, 1, 4)$ toward $P(1, 2, 3)$.

4.4 Find the amount of energy required to move a 6-C charge from the origin to $P(3, 1, -1)$ in the field $\mathbf{E} = 2x\mathbf{a}_x - 3y^2\mathbf{a}_y + 4\mathbf{a}_z$ V/m along the straight-line path $x = -3z,\ y = x + 2z$.

4.5 Compute the value of $\int_A^P \mathbf{G} \cdot d\mathbf{L}$ for $\mathbf{G} = 2y\mathbf{a}_x$ with $A(1, -1, 2)$ and $P(2, 1, 2)$ using the path: (*a*) straight-line segments $A(1, -1, 2)$ to $B(1, 1, 2)$ to $P(2, 1, 2)$; (*b*) straight-line segments $A(1, -1, 2)$ to $C(2, -1, 2)$ to $P(2, 1, 2)$.

4.6 Let $\mathbf{G} = 4x\mathbf{a}_x + 2z\mathbf{a}_y + 2y\mathbf{a}_z$. Given an initial point $P(2, 1, 1)$ and a final point $Q(4, 3, 1)$, find $\int \mathbf{G} \cdot d\mathbf{L}$ using the path: (*a*) straight line: $y = x - 1$, $z = 1$; (*b*) parabola: $6y = x^2 + 2$, $z = 1$.

4.7 Repeat Prob. 6 for $\mathbf{G} = 3xy^2\mathbf{a}_x + 2z\mathbf{a}_y$.

4.8 A point charge Q_1 is located at the origin in free space. Find the work done in carrying a charge Q_2 from: (*a*) $B(r_B, \theta_B, \phi_B)$ to $C(r_A, \theta_B, \phi_B)$ with θ and ϕ held constant; (*b*) $C(r_A, \theta_B, \phi_B)$ to $D(r_A, \theta_A, \phi_B)$ with r and ϕ held constant; (*c*) $D(r_A, \theta_A, \phi_B)$ to $A(r_A, \theta_A, \phi_A)$ with r and θ held constant.

4.9 A uniform surface charge density of 20 nC/m^2 is present on the spherical surface $r = 0.6$ cm in free space. (*a*) Find the absolute potential at $P(r = 1\text{ cm},\ \theta = 25^\circ,\ \phi = 50^\circ)$. (*b*) Find V_{AB}, given points $A(2\text{ cm},\ \theta = 30^\circ,\ \phi = 60^\circ)$ and $B(3\text{ cm}, 45^\circ, 90^\circ)$.

4.10 Given a surface charge density of 8 nC/m^2 on the plane $x = 2$, a line charge density of 30 nC/m on the line $x = 1$, $y = 2$, and a 1-μC point charge at $P(-1, -1, 2)$, find V_{AB} for points $A(3, 4, 0)$ and $B(4, 0, 1)$.

4.11 Let a uniform surface charge density of 5 nC/m^2 be present at the $z = 0$ plane, a uniform line charge density of 8 nC/m be located at $x = 0$, $z = 4$, and a point charge of 2 μC be present at $P(2, 0, 0)$. If $V = 0$ at $M(0, 0, 5)$, find V at $N(1, 2, 3)$.

4.12 Three point charges, 0.4 μC each, are located at $(0, 0, -1)$, $(0, 0, 0)$, and $(0, 0, 1)$, in free space. (*a*) Find an expression for the absolute potential as a function of z along the line $x = 0$, $y = 1$. (*b*) Sketch $V(z)$.

4.13 Three identical point charges of 4 pC each are located at the corners of an equilateral triangle 0.5 mm on a side in free space. How much work

must be done to move one charge to a point equidistant from the other two and on the line joining them?

4.14 Two 6-nC point charges are located at $(1, 0, 0)$ and $(-1, 0, 0)$ in free space. (*a*) Find V at $P(0, 0, z)$. (*b*) Find V_{max}. (*c*) Calculate $|dV/dz|$ on the z axis. (*d*) Find $|dV/dz|_{max}$.

4.15 Two uniform line charges, 8 nC/m each, are located at $x = 1$, $z = 2$, and at $x = -1$, $y = 2$, in free space. If the potential at the origin is 100 V, find V at $P(4, 1, 3)$.

4.16 Uniform surface charge distributions of 6, 4, and 2 nC/m^2 are present at $r = 2, 4$, and 6 cm, respectively, in free space. (*a*) Assume $V = 0$ at infinity, and find $V(r)$. (*b*) Calculate V at $r = 1, 3, 5$, and 7 cm. (*c*) Sketch V versus r for $1 < r < 10$ cm.

4.17 Uniform surface charge densities of 6 and 2 nC/m^2 are present at $\rho = 2$ and 6 cm, respectively, in free space. Assume $V = 0$ at $\rho = 4$ cm, and calculate V at $\rho =$ (*a*) 5 cm; (*b*) 7 cm.

4.18 The nonuniform linear charge density, $\rho_L = 8/(z^2 + 1)$ nC/m, lies along the z axis. Find the potential at $P(\rho = 1, 0, 0)$ in free space if $V = 0$ at $\rho = \infty$.

4.19 The annular surface, 1 cm $< \rho <$ 3 cm, $z = 0$, carries the nonuniform surface charge density $\rho_S = 5\rho$ nC/m^2. Find V at $P(0, 0, 2\text{ cm})$ if $V = 0$ at infinity.

4.20 Fig. 4.11 shows three separate charge distributions in the $z = 0$ plane in free space. (*a*) Find the total charge for each distribution. (*b*) Find the potential at $P(0, 0, 6)$ caused by each of the three charge distributions acting alone. (*c*) Find V_P.

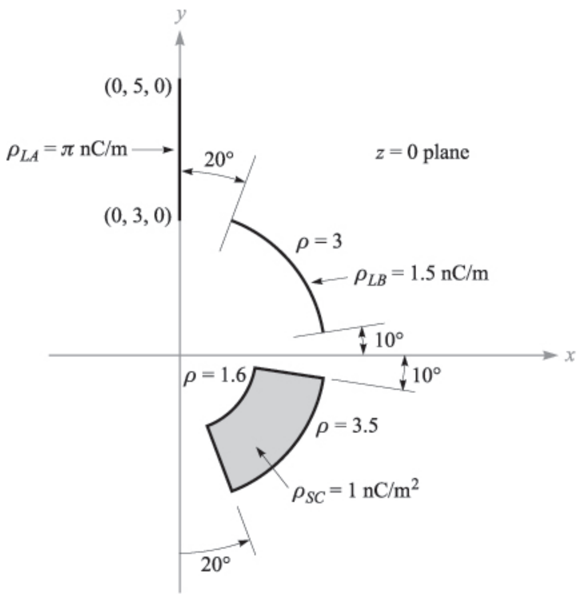


FIGURE 4.11
See Prob. 20.

4.21 Let $V = 2xy^2z^3 + 3\ln(x^2 + 2y^2 + 3z^2)$ V in free space. Evaluate each of the following quantities at $P(3, 2, -1)$: (*a*) V; (*b*) $|V|$; (*c*) $\mathbf{E}$; (*d*) $|\mathbf{E}|$; (*e*) $\mathbf{a}_N$; (*f*) $\mathbf{D}$.

4.22 It is known that the potential is given as $V = 80r^{0.6}$ V. Assuming free-space conditions, find: (*a*) $\mathbf{E}$; (*b*) the volume charge density at $r = 0.5$ m; (*c*) the total charge lying within the surface $r = 0.6$.

4.23 It is known that the potential is given as $V = 80\rho^{0.6}$ V. Assuming free-space conditions, find: (*a*) $\mathbf{E}$; (*b*) the volume charge density at $\rho = 0.5$ m; (*c*) the total charge lying within the closed surface $\rho = 0.6, 0 < z < 1$.

4.24 Given the potential field $V = 80r^2 \cos\theta$ and a point $P(2.5, \theta = 30^\circ, \phi = 60^\circ)$ in free space, find at P: (*a*) V; (*b*) $\mathbf{E}$; (*c*) $\mathbf{D}$; (*d*) ρ_v; (*e*) dV/dN; (*f*) $\mathbf{a}_N$

4.25 Within the cylinder $\rho = 2, 0 < z < 1$, the potential is given by $V = 100 + 50\rho + 150\rho \sin\phi$ V. (*a*) Find $V, \mathbf{E}, \mathbf{D}$, and ρ_v at $P(1, 60^\circ, 0.5)$ in free space. (*b*) How much charge lies within the cylinder?

4.26 A dipole having $Qd/(4\pi\epsilon_0) = 100\text{V} \cdot \text{m}^2$ is located at the origin in free space and aligned so that its moment is in the $\mathbf{a}_z$ direction. (*a*) Sketch $|V(r = 1, \theta, \phi = 0)|$ versus θ on polar graph paper (homemade if you wish). (*b*) Sketch $|\mathbf{E}(r = 1, \theta, \phi = 0)|$ versus θ on polar paper.

4.27 Two point charges, 1 nC at $(0, 0, 0.1)$ and -1 nC at $(0, 0, -0.1)$, are in free space. (*a*) Calculate V at $P(0.3, 0, 0.4)$, (*b*) Calculate $|\mathbf{E}|$ at P. (*c*) Now treat the two charges as a dipole at the origin and find V at P.

4.28 A dipole located at the origin in free space has a moment $\mathbf{p} = 2 \times 10^{-9}\mathbf{a}_z\,\text{C} \cdot \text{m}$. At what points on the line $y = z, x = 0$ is: (*a*) $|E_\theta| = 1$ mV/m? (*b*) $|E_r| = 1$ mV/m?

4.29 A dipole having a moment $\mathbf{p} = 3\mathbf{a}_x - 5\mathbf{a}_y + 10\mathbf{a}_z\,\text{nC} \cdot \text{m}$ is located at $Q(1, 2, -4)$ in free space. Find V at $P(2, 3, 4)$.

4.30 A dipole, having a moment of $\mathbf{p} = 2\mathbf{a}_z\,\text{nC} \cdot \text{m}$, is located at the origin in free space. Give the magnitude of $\mathbf{E}$ and its direction $\mathbf{a}_E$ in cartesian components at $r = 100$ m, $\phi = 90^\circ$, and $\theta =$: (*a*) 0°; (*b*) 30°; (*c*) 90°.

4.31 A potential field in free space is expressed as $V = 20/(xyz)\,V$. (*a*) Find the total energy stored within the cube $1 < x, y, z < 2$. (*b*) What value would be obtained by assuming a uniform energy density equal to the value at the center of the cube?

4.32 In the region of free space where $2 < r < 3, 0.4\pi < \theta < 0.6\pi, 0 < \phi < \pi/2$, let $\mathbf{E} = \dfrac{k}{r^2}\mathbf{a}_r$. (*a*) Find a positive value for k so that the total energy stored is exactly 1 J. (*b*) Show that the surface $\theta = 0.6\pi$ is an equipotential surface. (*c*) Find V_{AB} given points $A(2, \theta = \pi/2, \phi = \pi/3)$ and $B(3, \pi/2, \pi/4)$.

4.33 A copper sphere of radius 4 cm carries a uniformly distributed total charge of 5 µC on its surface in free space. (*a*) Use Gauss's law to find $\mathbf{D}$ external to the sphere. (*b*) Calculate the total energy stored in the electrostatic field. (*c*) Use $W_E = Q^2/(2C)$ to calculate the capacitance of the isolated sphere.

4.34 Given the potential field in free space, $V = 80\phi\mathbf{a}_\phi$ V (cyl. coord.), find: (*a*) the energy stored in the region $2 < \rho < 4$ cm, $0 < \phi < 0.2\pi$, $0 < z < 1$ m; (*b*) the potential difference V_{AB} for $A(3\text{ cm}, \phi = 0, z = 0)$ and $B(3\text{ cm}, 0.2\pi, 1\text{ m})$; (*c*) the maximum value of the energy density in the specified region.

4.35 Four 0.8-nC point charges are located in free space at the corners of a square 4 cm on a side. (*a*) Find the total potential energy stored. (*b*) A fifth 0.8-μC charge is installed at the center of the square. Again find the total stored energy.

CHAPTER
5

CONDUCTORS, DIELECTRICS, AND CAPACITANCE

In this chapter we intend to apply the laws and methods of the previous chapters to some of the materials with which an engineer must work. After defining current and current density and developing the fundamental continuity equation, we shall consider a conducting material and present Ohm's law in both its microscopic and macroscopic forms. With these results we may calculate resistance values for a few of the simpler geometrical forms that resistors may assume. Conditions which must be met at conductor boundaries are next obtained, and this knowledge enables us to introduce the use of images.

After a brief consideration of a general semiconductor, we shall investigate the polarization of dielectric materials and define relative permittivity, or the dielectric constant, an important engineering parameter. Having both conductors and dielectrics, we may then put them together to form capacitors. Most of the work of the previous chapters will be required to determine the capacitance of the several capacitors which we shall construct.

The fundamental electromagnetic principles on which resistors and capacitors depend are really the subject of this chapter; the inductor will not be introduced until Chap. 9.

5.1 CURRENT AND CURRENT DENSITY

Electric charges in motion constitute a *current*. The unit of current is the ampere (A), defined as a rate of movement of charge passing a given reference point (or crossing a given reference plane) of one coulomb per second. Current is symbolized by I, and therefore

$$I = \frac{dQ}{dt} \tag{1}$$

Current is thus defined as the motion of positive charges, even though conduction in metals takes place through the motion of electrons, as we shall see shortly.

In field theory we are usually interested in events occurring at a point rather than within some large region, and we shall find the concept of *current density*, measured in amperes per square meter (A/m^2), more useful. Current density is a vector[1] represented by **J**.

The increment of current ΔI crossing an incremental surface ΔS normal to the current density is

$$\Delta I = J_N \Delta S$$

and in the case where the current density is not perpendicular to the surface,

$$\Delta I = \mathbf{J} \cdot \Delta \mathbf{S}$$

Total current is obtained by integrating,

$$\boxed{I = \int_S \mathbf{J} \cdot d\mathbf{S}} \tag{2}$$

Current density may be related to the velocity of volume charge density at a point. Consider the element of charge $\Delta Q = \rho_v \Delta v = \rho_v \, \Delta S \, \Delta L$, as shown in Fig. 5.1*a*. To simplify the explanation, let us assume that the charge element is oriented with its edges parallel to the coordinate axes, and that it possesses only an x component of velocity. In the time interval Δt, the element of charge has moved a distance Δx, as indicated in Fig. 5.1*b*. We have therefore moved a charge $\Delta Q = \rho_v \, \Delta S \, \Delta x$ through a reference plane perpendicular to the direction of motion in a time increment Δt, and the resultant current is

$$\Delta I = \frac{\Delta Q}{\Delta t} = \rho_v \, \Delta S \frac{\Delta x}{\Delta t}$$

As we take the limit with respect to time, we have

$$\Delta I = \rho_v \, \Delta S \, v_x$$

[1] Current is not a vector, for it is easy to visualize a problem in which a total current I in a conductor of nonuniform cross section (such as a sphere) may have a different direction at each point of a given cross section. Current in an exceedingly fine wire, or a *filamentary current*, is occasionally defined as a vector, but we usually prefer to be consistent and give the direction to the filament, or path, and not to the current.

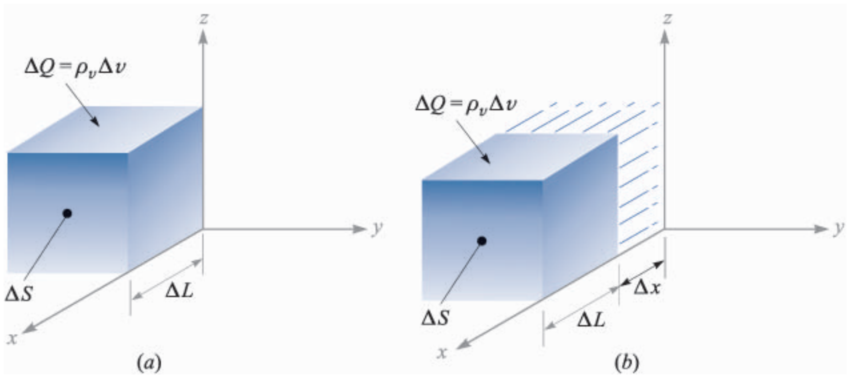


FIGURE 5.1
An increment of charge, $\Delta Q = \rho_v \Delta S\, \Delta L$, which moves a distance Δx in a time Δt, produces a component of current density in the limit of $J_x = \rho_v v_x$.

where v_x represents the x component of the velocity $\mathbf{v}$.[2] In terms of current density, we find

$$J_x = \rho_v\, v_x$$

and in general

$$\boxed{\mathbf{J} = \rho_v \mathbf{v}} \tag{3}$$

This last result shows very clearly that charge in motion constitutes a current. We call this type of current a *convention current*, and $\mathbf{J}$ or $\rho_v \mathbf{v}$ is the *convection current density*. Note that the convection current density is related linearly to charge density as well as to velocity. The mass rate of flow of cars (cars per square foot per second) in the Holland Tunnel could be increased either by raising the density of cars per cubic foot, or by going to higher speeds, if the drivers were capable of doing so.

✔ **D5.1.** Given the vector current density $\mathbf{J} = 10\rho^2 z \mathbf{a}_\rho - 4\rho\cos^2\phi\, \mathbf{a}_\phi$ A/m^2: (*a*) find the current density at $P(\rho = 3,\ \phi = 30^\circ,\ z = 2)$; (*b*) determine the total current flowing outward through the circular band $\rho = 3$, $0 < \phi < 2\pi$, $2 < z < 2.8$.

Ans. $180\mathbf{a}_\rho - 9\mathbf{a}_\phi$ A/m^2; 518 A

[2] The lowercase v is used both for volume and velocity. Note, however, that velocity always appears as a vector $\mathbf{v}$, a component v_x, or a magnitude $|\mathbf{v}|$, while volume appears only in differential form as dv or Δv.

5.2 CONTINUITY OF CURRENT

Although we are supposed to be studying static fields at this time, the introduction of the concept of current is logically followed by a discussion of the conservation of charge and the continuity equation. The principle of conservation of charge states simply that charges can be neither created nor destroyed, although equal amounts of positive and negative charge may be *simultaneously* created, obtained by separation, destroyed, or lost by recombination.

The continuity equation follows from this principle when we consider any region bounded by a closed surface. The current through the closed surface is

$$I = \oint_S \mathbf{J} \cdot d\mathbf{S}$$

and this *outward flow* of positive charge must be balanced by a decrease of positive charge (or perhaps an increase of negative charge) within the closed surface. If the charge inside the closed surface is denoted by Q_i, then the rate of decrease is $-dQ_i/dt$ and the principle of conservation of charge requires

$$I = \oint_S \mathbf{J} \cdot d\mathbf{S} = -\frac{dQ_i}{dt} \tag{4}$$

It might be well to answer here an often-asked question. "Isn't there a sign error? I thought $I = dQ/dt$." The presence or absence of a negative sign depends on what current and charge we consider. In circuit theory we usually associate the current flow *into* one terminal of a capacitor with the time rate of increase of charge on that plate. The current of (4), however, is an *outward-flowing* current.

Equation (4) is the integral form of the continuity equation, and the differential, or point, form is obtained by using the divergence theorem to change the surface integral into a volume integral:

$$\oint_S \mathbf{J} \cdot d\mathbf{S} = \int_{\text{vol}} (\nabla \cdot \mathbf{J})\, dv$$

We next represent the enclosed charge Q_i by the volume integral of the charge density,

$$\int_{\text{vol}} (\nabla \cdot \mathbf{J})\, dv = -\frac{d}{dt}\int_{\text{vol}} \rho_v \, dv$$

If we agree to keep the surface constant, the derivative becomes a partial derivative and may appear within the integral,

$$\int_{\text{vol}} (\nabla \cdot \mathbf{J})\, dv = \int_{\text{vol}} -\frac{\partial \rho_v}{\partial t}\, dv$$

Since the expression is true for any volume, however small, it is true for an incremental volume,

$$(\nabla \cdot \mathbf{J})\, \Delta v = -\frac{\partial \rho_v}{\partial t} \Delta v$$

from which we have our point form of the continuity equation,

$$(\nabla \cdot \mathbf{J}) = -\frac{\partial \rho_v}{\partial t} \tag{5}$$

Remembering the physical interpretation of divergence, this equation indicates that the current, or charge per second, diverging from a small volume per unit volume is equal to the time rate of decrease of charge per unit volume at every point.

As a numerical example illustrating some of the concepts from the last two sections, let us consider a current density that is directed radially outward and decreases exponentially with time,

$$\mathbf{J} = \frac{1}{r} e^{-t} \mathbf{a}_r \quad \text{A/m}^2$$

Selecting an instant of time $t = 1$ s, we may calculate the total outward current at $r = 5$ m:

$$I = J_r S = \left(\tfrac{1}{5} e^{-1}\right)(4\pi 5^2) = 23.1 \quad \text{A}$$

At the same instant, but for a slightly larger radius, $r = 6$ m, we have

$$I = J_r S = \left(\tfrac{1}{6} e^{-1}\right)\left(4\pi 6^2\right) = 27.7 \quad \text{A}$$

Thus, the total current is larger at $r = 6$ than it is at $r = 5$.

To see why this happens, we need to look at the volume charge density and the velocity. We use the continuity equation first:

$$-\frac{\partial \rho_v}{\partial t} = \nabla \cdot \mathbf{J} = \nabla \cdot \left(\frac{1}{r} e^{-t} \mathbf{a}_r\right) = \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{1}{r} e^{-t}\right) = \frac{1}{r^2} e^{-t}$$

We next seek the volume charge density by integrating with respect to t. Since ρ_v is given by a partial derivative with respect to time, the "constant" of integration may be a function of r:

$$\rho_v = -\int \frac{1}{r^2} e^{-t}\, dt + K(r) = \frac{1}{r^2} e^{-t} + K(r)$$

If we assume that $\rho_v \to 0$ as $t \to \infty$, then $K(r) = 0$, and

$$\rho_v = \frac{1}{r^2} e^{-t} \quad \text{C/m}^3$$

We may now use $\mathbf{J} = \rho_v \mathbf{v}$ to find the velocity,

$$v_r = \frac{J_r}{\rho_v} = \frac{\frac{1}{r} e^{-t}}{\frac{1}{r^2} e^{-t}} = r \quad \text{m/s}$$

The velocity is greater at $r = 6$ than it is at $r = 5$, and we see that some (unspecified) force is accelerating the charge density in an outward direction.

In summary, we have a current density that is inversely proportional to r, a charge density that is inversely proportional to r^2, and a velocity and total current that are proportional to r. All quantities vary as e^{-t}.

✔ **D5.2.** Current density is given in cylindrical coordinates as $\mathbf{J} = -10^6 z^{1.5} \mathbf{a}_z$ A/m² in the region $0 \le \rho \le 20\,\mu$m; for $\rho \ge 20\,\mu$m, $\mathbf{J} = 0$. (*a*) Find the total current crossing the surface $z = 0.1$ m in the $\mathbf{a}_z$ direction. (*b*) If the charge velocity is 2×10^6 m/s at $z = 0.1$ m, find ρ_v there. (*c*) If the volume charge density at $z = 0.15$ m is $-2000\,\text{C/m}^3$, find the charge velocity there.

Ans. -39.7 mA; $-15.81\,\text{kC/m}^3$; -2900 m/s

5.3 METALLIC CONDUCTORS

Physicists today describe the behavior of the electrons surrounding the positive atomic nucleus in terms of the total energy of the electron with respect to a zero reference level for an electron at an infinite distance from the nucleus. The total energy is the sum of the kinetic and potential energies, and since energy must be given to an electron to pull it away from the nucleus, the energy of every electron in the atom is a negative quantity. Even though the picture has some limitations, it is convenient to associate these energy values with orbits surrounding the nucleus, the more negative energies corresponding to orbits of smaller radius. According to the quantum theory, only certain discrete energy levels, or energy states, are permissible in a given atom, and an electron must therefore absorb or emit discrete amounts of energy, or quanta, in passing from one level to another. A normal atom at absolute zero temperature has an electron occupying every one of the lower energy shells, starting outward from the nucleus and continuing until the supply of electrons is exhausted.

In a crystalline solid, such as a metal or a diamond, atoms are packed closely together, many more electrons are present, and many more permissible energy levels are available because of the interaction forces between adjacent atoms. We find that the energies which may be possessed by electrons are grouped into broad ranges, or "bands," each band consisting of very numerous, closely spaced, discrete levels. At a temperature of absolute zero, the normal solid also has every level occupied, starting with the lowest and proceeding in order until all the electrons are located. The electrons with the highest (least negative) energy levels, the valence electrons, are located in the *valence band*. If there are permissible higher-energy levels in the valence band, or if the valence band merges smoothly into a *conduction band*, then additional kinetic energy may be given to the valence electrons by an external field, resulting in an electron flow. The solid is called a *metallic conductor*. The filled valence band and the unfilled conduction band for a conductor at 0 K are suggested by the sketch in Fig. 5.2*a*.

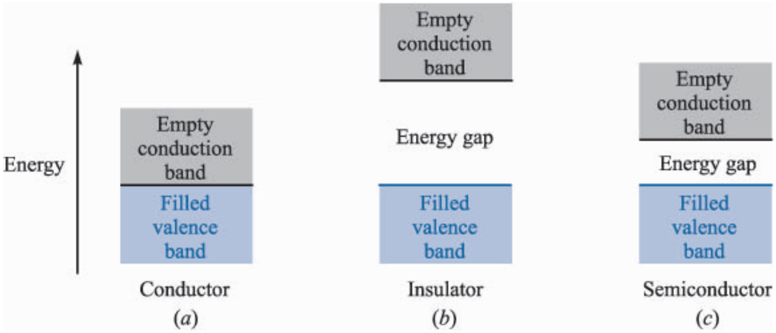


FIGURE 5.2
The energy-band structure in three different types of materials at 0 K. (*a*) The conductor exhibits no energy gap between the valence and conduction bands. (*b*) The insulator shows a large energy gap. (*c*) The semiconductor has only a small energy gap.

If, however, the electron with the greatest energy occupies the top level in the valence band and a gap exists between the valence band and the conduction band, then the electron cannot accept additional energy in small amounts, and the material is an insulator. This band structure is indicated in Fig. 5.2*b*. Note that if a relatively large amount of energy can be transferred to the electron, it may be sufficiently excited to jump the gap into the next band where conduction can occur easily. Here the insulator breaks down.

An intermediate condition occurs when only a small "forbidden region" separates the two bands, as illustrated by Fig. 5.2*c*. Small amounts of energy in the form of heat, light, or an electric field may raise the energy of the electrons at the top of the filled band and provide the basis for conduction. These materials are insulators which display many of the properties of conductors and are called *semiconductors*.

Let us first consider the conductor. Here the valence electrons, or *conduction*, or *free*, electrons, move under the influence of an electric field. With a field $\mathbf{E}$, an electron having a charge $Q = -e$ will experience a force

$$\mathbf{F} = -e\mathbf{E}$$

In free space the electron would accelerate and continuously increase its velocity (and energy); in the crystalline material the progress of the electron is impeded by continual collisions with the thermally excited crystalline lattice structure, and a constant average velocity is soon attained. This velocity $\mathbf{v}_d$ is termed the *drift velocity*, and it is linearly related to the electric field intensity by the *mobility* of the electron in the given material. We designate mobility by the symbol μ (mu), so that

$$\mathbf{v}_d = -\mu_e \mathbf{E} \tag{6}$$

where μ_ϵ is the mobility of an electron and is positive by definition. Note that the electron velocity is in a direction opposite to the direction of $\mathbf{E}$. Equation (6) also

shows that mobility is measured in the units of square meters per volt-second; typical values[3] are 0.0012 for aluminum, 0.0032 for copper, and 0.0056 for silver.

For these good conductors a drift velocity of a few inches per second is sufficient to produce a noticeable temperature rise and can cause the wire to melt if the heat cannot be quickly removed by thermal conduction or radiation.

Substituting (6) into Eq. (3) of Sec. 5.1, we obtain

$$\boxed{\mathbf{J} = -\rho_e \mu_e \mathbf{E}} \tag{7}$$

where ρ_e is the free-electron charge density, a negative value. The total charge density ρ_v is zero, since equal positive and negative charge is present in the neutral material. The negative value of ρ_e and the minus sign lead to a current density **J** that is in the same direction as the electric field intensity **E**.

The relationship between **J** and **E** for a metallic conductor, however, is also specified by the conductivity σ (sigma),

$$\boxed{\mathbf{J} = \sigma \mathbf{E}} \tag{8}$$

where σ is measured is siemens[4] per meter (S/m). One siemens (1 S) is the basic unit of conductance in the SI system, and is defined as one ampere per volt. Formerly, the unit of conductance was called the mho and symbolized by an *inverted* Ω. Just as the siemens honors the Siemens brothers, the reciprocal unit of resistance which we call the ohm (1 Ω is one volt per ampere) honors Georg Simon Ohm, a German physicist who first described the current-voltage relationship implied by (8). We call this equation the *point form of Ohm's law*; we shall look at the more common form of Ohm's law shortly.

First, however, it is informative to note the conductivity of several metallic conductors; typical values (in siemens per meter) are 3.82×10^7 for aluminum, 5.80×10^7 for copper, and 6.17×10^7 for silver. Data for other conductors may be found in Appendix C. On seeing data such as these, it is only natural to assume that we are being presented with *constant* values; this is essentially true. Metallic conductors obey Ohm's law quite faithfully, and it is a *linear* relationship; the conductivity is constant over wide ranges of current density and electric field intensity. Ohm's law and the metallic conductors are also described as *isotropic*, or having the same properties in every direction. A material which is not isotropic is called *anisotropic*, and we shall mention such a material a few pages from now.

---

[3] Wert and Thomson, p. 238, listed in the Suggested References at the end of this chapter.

[4] This is the family name of two German-born brothers, Karl Wilhelm and Werner von Siemens, who were famous engineer-inventors in the nineteenth century. Karl became a British subject and was knighted, becoming Sir William Siemens.

The conductivity is a function of temperature, however. The resistivity, which is the reciprocal of the conductivity, varies almost linearly with temperature in the region of room temperature, and for aluminum, copper, and silver it increases about 0.4 percent for a 1 K rise in temperature.[5] For several metals the resistivity drops abruptly to zero at a temperature of a few kelvin; this property is termed *superconductivity*. Copper and silver are not superconductors, although aluminum is (for temperatures below 1.14 K).

If we now combine (7) and (8), the conductivity may be expressed in terms of the charge density and the electron mobility,

$$\boxed{\sigma = -\rho_e \mu_e} \tag{9}$$

From the definition of mobility (6), it is now satisfying to note that a higher temperature infers a greater crystalline lattice vibration, more impeded electron progress for a given electric field strength, lower drift velocity, lower mobility, lower conductivity from (9), and higher resistivity as stated.

The application of Ohm's law in point form to a macroscopic (visible to the naked eye) region leads to a more familiar form. Initially, let us assume that **J** and **E** are *uniform*, as they are in the cylindrical region shown in Fig. 5.3. Since they are uniform,

$$I = \int_S \mathbf{J} \cdot d\mathbf{S} = JS \tag{10}$$

and

$$V_{ab} = -\int_b^a \mathbf{E} \cdot d\mathbf{L} = -\mathbf{E} \cdot \int_b^a d\mathbf{L} = -\mathbf{E} \cdot \mathbf{L}_{ba}$$

$$= \mathbf{E} \cdot \mathbf{L}_{ab} \tag{11}$$

or

$$V = EL$$

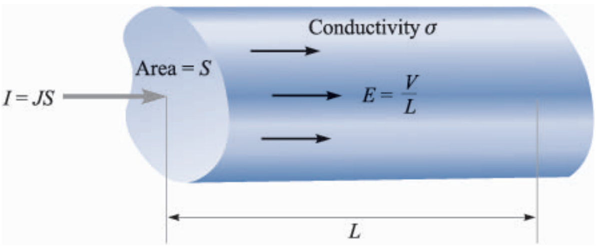


FIGURE 5.3 Uniform current density **J** and electric field intensity **E** in a cylindrical region of length L and cross-sectional area S. Here $V = IR$, where $R = L/\sigma S$.

[5] Copious temperature data for conducting materials are available in the "Standard Handbook for Electrical Engineers," listed among the Suggested References at the end of this chapter.

Thus

$$J = \frac{I}{S} = \sigma E = \sigma \frac{V}{L}$$

or

$$V = \frac{L}{\sigma S} I$$

The ratio of the potential difference between the two ends of the cylinder to the current entering the more positive end, however, is recognized from elementary circuit theory as the *resistance* of the cylinder, and therefore

$$\boxed{V = IR} \tag{12}$$

where

$$\boxed{R = \frac{L}{\sigma S}} \tag{13}$$

Equation (12) is, of course, known as *Ohm's law*, and (13) enables us to compute the resistance R, measured in ohms (abbreviated as Ω), of conducting objects which possess uniform fields. If the fields are not uniform, the resistance may still be defined as the ratio of V to I, where V is the potential difference between two specified equipotential surfaces in the material and I is the total current crossing the more positive surface into the material. From the general integral relationships in (10) and (11), and from Ohm's law (8), we may write this general expression for resistance when the fields are nonuniform,

$$R = \frac{V_{ab}}{I} = \frac{-\int_b^a \mathbf{E} \cdot d\mathbf{L}}{\int_S \sigma \mathbf{E} \cdot d\mathbf{S}} \tag{14}$$

The line integral is taken between two equipotential surfaces in the conductor, and the surface integral is evaluated over the more positive of these two equipotentials. We cannot solve these nonuniform problems at this time, but we should be able to solve several of them after perusing Chaps. 6 and 7.

Example 5.1

As an example of the determination of the resistance of a cylinder, let us find the resistance of a 1-mile length of #16 copper wire, which has a diameter of 0.0508 in.

Solution. The diameter of the wire is $0.0508 \times 0.0254 = 1.291 \times 10^{-3}$ m, the area of the cross section is $\pi(1.291 \times 10^{-3}/2)^2 = 1.308 \times 10^{-6}\ \text{m}^2$, and the length is 1609 m. Using a conductivity of 5.80×10^7 S/m, the resistance of the wire is therefore

$$R = \frac{1609}{(5.80 \times 10^7)(1.308 \times 10^{-6})} = 21.2 \quad \Omega$$

This wire can safely carry about 10 A dc, corresponding to a current density of $10/(1.308 \times 10^{-6}) = 7.65 \times 10^6 \ \text{A/m}^2$, or $7.65 \ \text{A/mm}^2$. With this current the potential difference between the two ends of the wire is 212 V, the electric field intensity is 0.312 V/m, the drift velocity is 0.000 422 m/s, or a little more than one furlong a week, and the free-electron charge density is $-1.81 \times 10^{10} \ \text{C/m}^3$, or about one electron in a cube two angstroms on a side.

✔ **D5.3.** Find the magnitude of the current density in a sample of silver for which $\sigma = 6.17 \times 10^7$ S/m and $\mu_e = 0.0056 \ \text{m}^2/\text{V} \cdot \text{s}$ if: (*a*) the drift velocity is 1.5 μm/*s*; (*b*) the electric field intensity is 1 mV/m; (*c*) the sample is a cube 2.5 mm on a side having a voltage of 0.4 mV between opposite faces; (*d*) the sample is a cube 2.5 mm on a side carrying a total current of 0.5 A.

Ans. 16.53 kA/m^2; 61.7 kA/m^2; 9.87 MA/m^2; 80.0 kA/m^2

✔ **D5.4.** A copper conductor has a diameter of 0.6 in and it is 1200 ft long. Assume that it carries a total dc current of 50 A. (*a*) Find the total resistance of the conductor. (*b*) What current density exists in it? (*c*) What is the dc voltage between the conductor ends? (*d*) How much power is dissipated in the wire?

Ans. 0.0346 Ω; 2.74×10^5 A/m^2; 1.729 V; 86.4 W

5.4 CONDUCTOR PROPERTIES AND BOUNDARY CONDITIONS

Once again we must temporarily depart from our assumed static conditions and let time vary for a few microseconds to see what happens when the charge distribution is suddenly unbalanced within a conducting material. Let us suppose, for the sake of the argument, that there suddenly appear a number of electrons in the interior of a conductor. The electric fields set up by these electrons are not counteracted by any positive charges, and the electrons therefore begin to accelerate away from each other. This continues until the electrons reach the surface of the conductor or until a number of electrons equal to the number injected have reached the surface.

Here the outward progress of the electrons is stopped, for the material surrounding the conductor is an insulator not possessing a convenient conduction band. No charge may remain within the conductor. If it did, the resulting electric field would force the charges to the surface.

Hence the final result within a conductor is zero charge density, and a surface charge density resides on the exterior surface. This is one of the two characteristics of a good conductor.

The other characteristic, stated for static conditions in which no current may flow, follows directly from Ohm's law: the electric field intensity within the conductor is zero. Physically, we see that if an electric field were present, the

conduction electrons would move and produce a current, thus leading to a nonstatic condition.

Summarizing for electrostatics, no charge and no electric field may exist at any point *within* a conducting material. Charge may, however, appear on the surface as a surface charge density, and our next investigation concerns the fields *external* to the conductor.

We wish to relate these external fields to the charge on the surface of the conductor. The problem is a simple one, and we may first talk our way to the solution with little mathematics.

If the external electric field intensity is decomposed into two components, one tangential and one normal to the conductor surface, the tangential component is seen to be zero. If it were not zero, a tangential force would be applied to the elements of the surface charge, resulting in their motion and nonstatic conditions. Since static conditions are assumed, the tangential electric field intensity and electric flux density are zero.

Gauss's law answers our questions concerning the normal component. The electric flux leaving a small increment of surface must be equal to the charge residing on that incremental surface. The flux cannot penetrate into the conductor, for the total field there is zero. It must then leave the surface normally. Quantitatively, we may say that the electric flux density in coulombs per square meter leaving the surface normally is equal to the surface charge density in coulombs per square meter, or $D_N = \rho_S$.

If we use some of our previously derived results in making a more careful analysis (and incidentally introducing a general method which we must use later), we should set up a conductor-free space boundary (Fig. 5.4) showing tangential and normal components of **D** and **E** on the free-space side of the boundary. Both fields are zero in the conductor. The tangential field may be determined by applying Sec. 4.5, Eq. (21),

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0$$

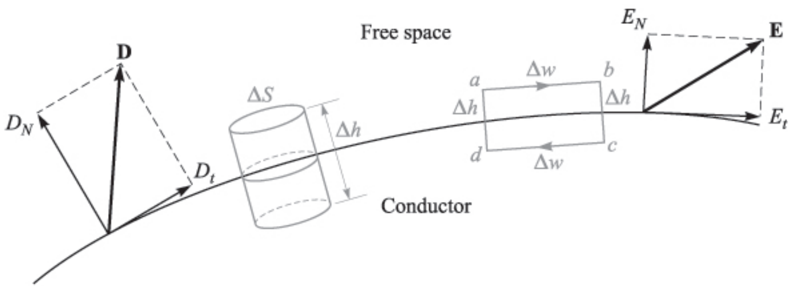


FIGURE 5.4
An appropriate closed path and gaussian surface are used to determine boundary conditions at a conductor-free space boundary; $E_t = 0$ and $D_N = \rho_S$.

around the small closed path *abcda*. The integral must be broken up into four parts

$$\int_a^b + \int_b^c + \int_c^d + \int_d^a = 0$$

Remembering that $\mathbf{E} = 0$ within the conductor, we let the length from a to b or c to d be Δw and from b to c or d to a be Δh, and obtain

$$E_t \Delta w - E_{N,\text{at}\,b} \tfrac{1}{2} \Delta h + E_{N,\text{at}\,a} \tfrac{1}{2} \Delta h = 0$$

As we allow Δh to approach zero, keeping Δw small but finite, it makes no difference whether or not the normal fields are equal at a and b, for Δh causes these products to become negligibly small. Hence

$$E_t \Delta w = 0$$

and therefore

$$E_t = 0$$

The condition on the normal field is found most readily by considering D_N rather than E_N and choosing a small cylinder as the gaussian surface. Let the height be Δh and the area of the top and bottom faces be ΔS. Again we shall let Δh approach zero. Using Gauss's law,

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q$$

we integrate over the three distinct surfaces

$$\int_{\text{top}} + \int_{\text{bottom}} + \int_{\text{sides}} = Q$$

and find that the last two are zero (for different reasons). Then

$$D_N \Delta S = Q = \rho_S \Delta S$$

or

$$D_N = \rho_S$$

These are the desired *boundary conditions* for the conductor-free space boundary in electrostatics,

$$\boxed{D_t = E_t = 0} \tag{15}$$

$$\boxed{D_N = \epsilon_0 E_N = \rho_S} \tag{16}$$

The electric flux leaves the conductor in a direction normal to the surface, and the value of the electric flux density is numerically equal to the surface charge density.

An immediate and important consequence of a zero tangential electric field intensity is the fact that a conductor surface is an equipotential surface. The evaluation of the potential difference between any two points on the surface by the line integral leads to a zero result, because the path may be chosen on the surface itself where $\mathbf{E} \cdot d\mathbf{L} = 0$.

To summarize the principles which apply to conductors in electrostatic fields, we may state that

1. The static electric field intensity inside a conductor is zero.
2. The static electric field intensity at the surface of a conductor is everywhere directed normal to that surface.
3. The conductor surface is an equipotential surface.

Using these three principles, there are a number of quantities that may be calculated at a conductor boundary, given a knowledge of the potential field.

Example 5.2

Given the potential,

$$V = 100(x^2 - y^2)$$

and a point $P(2, -1, 3)$ that is stipulated to lie on a conductor-free space boundary, let us find V, $\mathbf{E}$, $\mathbf{D}$, and ρ_S at P, and also the equation of the conductor surface.

Solution. The potential at point P is

$$V_P = 100[2^2 - (-1)^2] = 300 \quad \text{V}$$

Since the conductor is an equipotential surface, the potential at the entire surface must be 300 V. Moreover, if the conductor is a solid object, then the potential everywhere in and on the conductor is 300 V, for $\mathbf{E} = 0$ within the conductor.

The equation representing the locus of all points having a potential of 300 V is

$$300 = 100(x^2 - y^2)$$

or

$$x^2 - y^2 = 3$$

This is therefore the equation of the conductor surface; it happens to be a hyperbolic cylinder, as shown in Fig. 5.5. Let us assume arbitrarily that the solid conductor lies above and to the right of the equipotential surface at point P, while free space is down and to the left.

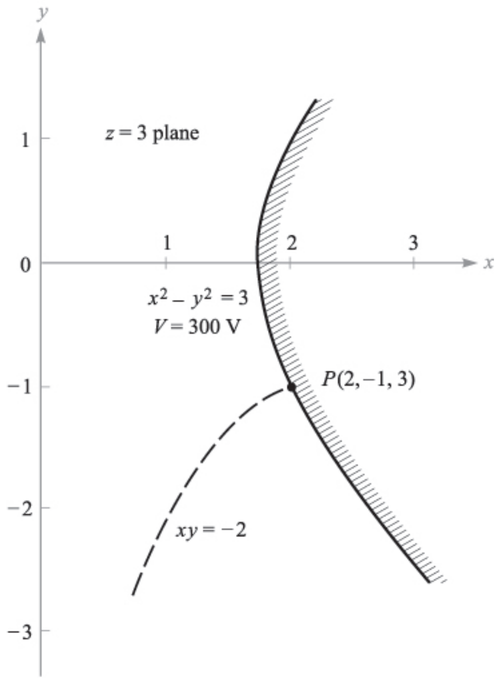


FIGURE 5.5
Given point $P(2, -1, 3)$ and the potential field, $V = 100(x^2 - y^2)$, we find the equipotential surface through P is $x^2 - y^2 = 3$, and the streamline through P is $xy = -2$.

Next, we find **E** by the gradient operation,

$$\mathbf{E} = -100\nabla(x^2 - y^2) = -200x\mathbf{a}_x + 200y\mathbf{a}_y$$

At point P,

$$\mathbf{E}_p = -400\mathbf{a}_x - 200\mathbf{a}_y \quad \text{V/m}$$

Since $\mathbf{D} = \epsilon_0\mathbf{E}$, we have

$$\mathbf{D}_P = 8.854 \times 10^{-12}\mathbf{E}_P = -3.54\mathbf{a}_x - 1.771\mathbf{a}_y \quad \text{nC/m}^2$$

The field is directed downward and to the left at P; it is normal to the equipotential surface. Therefore,

$$D_N = |\mathbf{D}_P| = 3.96 \quad \text{nC/m}^2$$

Thus, the surface charge density at P is

$$\rho_{S,P} = D_N = 3.96 \quad \text{nC/m}^2$$

Note that if we had taken the region to the left of the equipotential surface as the conductor, the **E** field would *terminate* on the surface charge and we would let $\rho_S = -3.96\,\text{nC/m}^2$.

Example 5.3

Finally, let us determine the equation of the streamline passing through P.

Solution. We see that

$$\frac{E_y}{E_x} = \frac{200y}{-200x} = -\frac{y}{x} = \frac{dy}{dx}$$

Thus,

$$\frac{dy}{y} + \frac{dx}{x} = 0$$

and

$$\ln y + \ln x = C_1$$

Therefore,

$$xy = C_2$$

The line (or surface) through P is obtained when $C_2 = (2)(-1) = -2$. Thus, the streamline is the trace of another hyperbolic cylinder,

$$xy = -2$$

This is also shown on Fig. 5.5.

✔ **D5.5.** Given the potential field in free space, $V = 100 \sinh 5x \sin 5y$ V, and a point $P(0.1, 0.2, 0.3)$, find at P: (*a*) V; (*b*) $\mathbf{E}$; (*c*) $|\mathbf{E}|$; (*d*) $|\rho_S|$ if it is known that P lies on a conductor surface.

Ans. 43.8 V; $-474\mathbf{a}_x - 140.8\mathbf{a}_y$ V/m; 495 V/m; 4.38 nC/m^2

5.5 THE METHOD OF IMAGES

One important characteristic of the dipole field that we developed in the last chapter is the infinite plane at zero potential that exists midway between the two charges. Such a plane may be represented by a vanishingly thin conducting plane that is infinite in extent. The conductor is an equipotential surface at a potential $V = 0$, and the electric field intensity is therefore normal to the surface. Thus, if we replace the dipole configuration shown in Fig. 5.6*a* with the single charge and conducting plane shown in Fig. 5.6*b*, the fields in the upper half of each figure are the same. Below the conducting plane, all fields are zero since we have not provided any charges in that region. Of course, we might also substitute a single negative charge below a conducting plane for the dipole arrangement and obtain equivalence for the fields in the lower half of each region.

If we approach this equivalence from the opposite point of view, we begin with a single charge above a perfectly conducting plane and then see that we may maintain the same fields above the plane by removing the plane and locating a negative charge at a symmetrical location below the plane. This charge is called the *image* of the original charge, and it is the negative of that value.

If we can do this once, linearity allows us to do it again and again, and thus *any* charge configuration above an infinite ground plane may be replaced by an

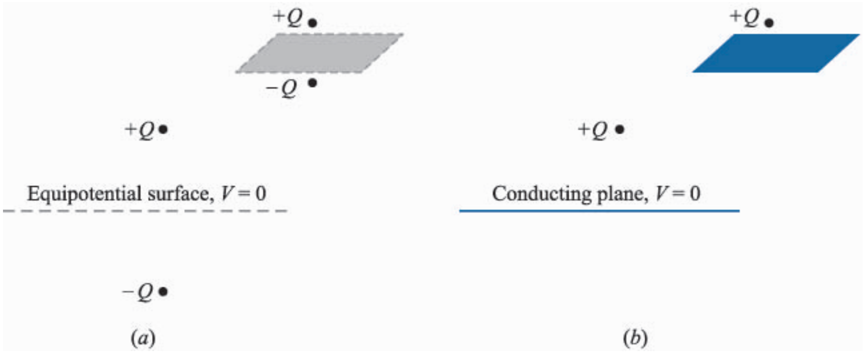


FIGURE 5.6
(a) Two equal but opposite charges may be replaced by (b) a single charge and a conducting plane without affecting the fields above the $V = 0$ surface.

arrangement composed of the given charge configuration, its image, and no conducting plane. This is suggested by the two illustrations of Fig. 5.7. In many cases, the potential field of the new system is much easier to find since it does not contain the conducting plane with its unknown surface charge distribution.

As an example of the use of images, let us find the surface charge density at $P(2, 5, 0)$ on the conducting plane $z = 0$ if there is a line charge of 30 nC/m located at $x = 0$, $z = 3$, as shown in Fig. 5.8a. We remove the plane and install an image line charge of −30 nC/m at $x = 0$, $z = -3$, as illustrated in Fig. 5.8b. The field at P may now be obtained by superposition of the known fields of the line charges. The radial vector from the positive line charge to P is $\mathbf{R}_+ = 2\mathbf{a}_x - 3\mathbf{a}_z$, while $\mathbf{R}_- = 2\mathbf{a}_x + 3\mathbf{a}_z$. Thus, the individual fields are

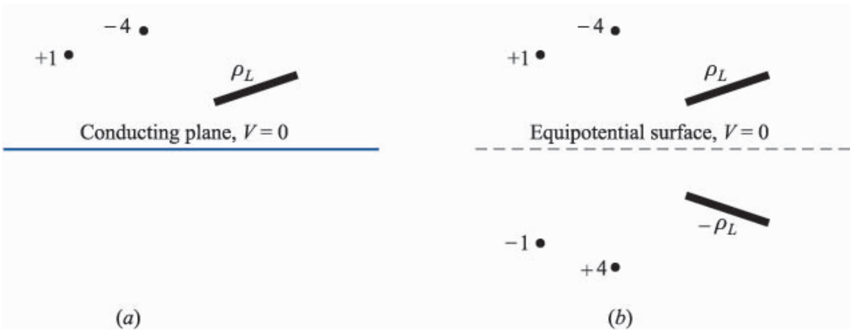


FIGURE 5.7
(a) A given charge configuration above an infinite conducting plane may be replaced by (b) the given charge configuration plus the image configuration, without the conducting plane.

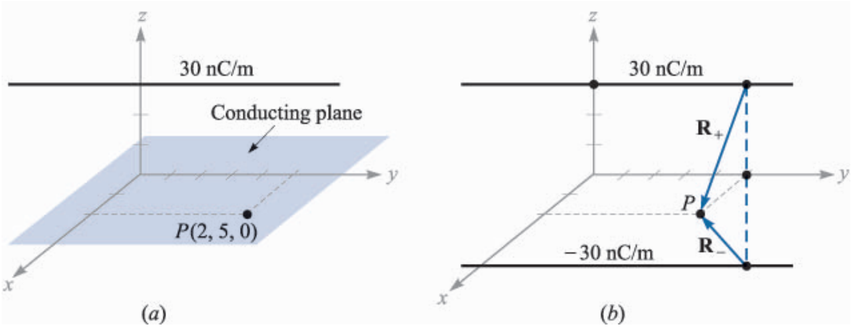


FIGURE 5.8
(*a*) A line charge above a conducting plane. (*b*) The conductor is removed, and the image of the line charge is added.

$$\mathbf{E}_+ = \frac{\rho_L}{2\pi\epsilon_0 R_+}\mathbf{a}_{R+} = \frac{30\times 10^{-9}}{2\pi\epsilon_0\sqrt{13}}\frac{2\mathbf{a}_x - 3\mathbf{a}_z}{\sqrt{13}}$$

and
$$\mathbf{E}_- = \frac{30\times 10^{-9}}{2\pi\epsilon_0\sqrt{13}}\frac{2\mathbf{a}_x + 3\mathbf{a}_z}{\sqrt{13}}$$

Adding these results, we have

$$\mathbf{E} = \frac{-180\times 10^{-9}\mathbf{a}_z}{2\pi\epsilon_0(13)} = -249\mathbf{a}_z \quad \text{V/m}$$

This then is the field at (or just above) P in both the configurations of Fig. 5.8, and it is certainly satisfying to note that the field is normal to the conducting plane, as it must be. Thus, $\mathbf{D} = \epsilon_0\mathbf{E} = -2.20\mathbf{a}_z\ \text{nC/m}^2$, and since this is directed *toward* the conducting plane, ρ_S is negative and has a value of $-2.20\ \text{nC/m}^2$ at P.

✔ **D5.6.** A perfectly conducting plane is located in free space at $x = 4$, and a uniform infinite line charge of 40 nC/m lies along the line $x = 6$, $y = 3$. Let $V = 0$ at the conducting plane. At $P(7, -1, 5)$ find: (*a*) V; (*b*) $\mathbf{E}$.

Ans. −316 V; $-45.4\mathbf{a}_x$ V/m.

5.6 SEMICONDUCTORS

If we now turn our attention to an intrinsic semiconductor material, such as pure germanium or silicon, two types of current carriers are present, electrons and *holes*. The electrons are those from the top of the filled valence band which have received sufficient energy (usually thermal) to cross the relatively small forbidden band into the conduction band. The forbidden-band energy gap in typical semiconductors is of the order of one electronvolt. The vacancies left by these electrons represent unfilled energy states in the valence band which may also move

from atom to atom in the crystal. The vacancy is called a *hole*, and many semiconductor properties may be described by treating the hole as if it had a positive charge of e, a mobility, μ_h, and an effective mass comparable to that of the electron. Both carriers move in an electric field, and they move in opposite directions; hence each contributes a component of the total current which is in the same direction as that provided by the other. The conductivity is therefore a function of both hole and electron concentrations and mobilities,

$$\boxed{\sigma = -\rho_e \mu_e + \rho_h \mu_h} \tag{17}$$

For pure, or *intrinsic*, silicon the electron and hole mobilities are 0.12 and 0.025, respectively, while for germanium, the mobilities are, respectively, 0.36 and 0.17. These values are given in square meters per volt-second and range from 10 to 100 times as large as those for aluminum, copper, silver, and other metallic conductors.[6] The mobilities listed above are given for a temperature of 300 K.

The electron and hole concentrations depend strongly on temperature. At 300 K the electron and hole volume charge densities are both 0.0024 C/m^3 in magnitude in intrinsic silicon and 3.0 C/m^3 in intrinsic germanium. These values lead to conductivities of 0.000 35 S/m in silicon and 1.6 S/m in germanium. As temperature increases, the mobilities decrease, but the charge densities increase very rapidly. As a result, the conductivity of silicon increases by a factor of 10 as the temperature increases from 300 to about 330 K and decreases by a factor of 10 as the temperature drops from 300 to about 275 K. Note that the conductivity of the intrinsic semiconductor increases with temperature, while that of a metallic conductor decreases with temperature; this is one of the characteristic differences between the metallic conductors and the intrinsic semiconductors.

Intrinsic semiconductors also satisfy the point form of Ohm's law; that is, the conductivity is reasonably constant with current density and with the direction of the current density.

The number of charge carriers and the conductivity may both be increased dramatically by adding very small amounts of impurities. *Donor* materials provide additional electrons and form *n-type* semiconductors, while *acceptors* furnish extra holes and form *p-type* materials. The process is known as *doping*, and a donor concentration in silicon as low as one part in 10^7 causes an increase in conductivity by a factor of 10^5.

The range of value of the conductivity is extreme as we go from the best insulating materials to semiconductors and the finest conductors. In siemens per meter, σ ranges from 10^{-17} for fused quartz, 10^{-7} for poor plastic insulators, and roughly unity for semiconductors to almost 10^8 for metallic conductors at room temperature. These values cover the remarkably large range of some 25 orders of magnitude.

---

[6] Mobility values for semiconductors are given in Refs. 2, 4, and 7 listed at the end of this chapter.

✔ **D5.7.** Using the values given in this section for the electron and hole mobilities in silicon at 300 K, and assuming hole and electron charge densities are 0.0029 C/m^3 and $-0.0029\ C/m^3$, respectively, find: (*a*) the component of the conductivity due to holes; (*b*) the component of the conductivity due to electrons; (*c*) the conductivity.

Ans. 0.0725 S/m; 0.348 S/m; 0.421 S/m

5.7 THE NATURE OF DIELECTRIC MATERIALS

Although we have mentioned insulators and dielectric materials, we do not as yet have any quantitative relationships in which they are involved. We shall soon see, however, that a dielectric in an electric field can be viewed as a free-space arrangement of microscopic electric dipoles which are composed of positive and negative charges whose centers do not quite coincide.

These are not free charges, and they cannot contribute to the conduction process. Rather, they are bound in place by atomic and molecular forces and can only shift positions slightly in response to external fields. They are called *bound* charges, in contrast to the free charges that determine conductivity. The bound charges can be treated as any other sources of the electrostatic field. If we did not wish to, therefore, we would not need to introduce the dielectric constant as a new parameter or to deal with permittivities different from the permittivity of free space; however, the alternative would be to consider *every charge within a piece of dielectric material*. This is too great a price to pay for using all our previous equations in an unmodified form, and we shall therefore spend some time theorizing about dielectrics in a qualitative way; introducing polarization **P**, permittivity ϵ, and relative permittivity ϵ_R; and developing some quantitative relationships involving these new quantities.

The characteristic which all dielectric materials have in common, whether they are solid, liquid, or gas, and whether or not they are crystalline in nature, is their ability to store electric energy. This storage takes place by means of a shift in the relative positions of the internal, bound positive and negative charges against the normal molecular and atomic forces.

This displacement against a restraining force is analogous to lifting a weight or stretching a spring and represents potential energy. The source of the energy is the external field, the motion of the shifting charges resulting perhaps in a transient current through a battery which is producing the field.

The actual mechanism of the charge displacement differs in the various dielectric materials. Some molecules, termed *polar* molecules, have a permanent displacement existing between the centers of "gravity" of the positive and negative charges, and each pair of charges acts as a dipole. Normally the dipoles are oriented in a random way throughout the interior of the material, and the action of the external field is to align these molecules, to some extent, in the same direction. A sufficiently strong field may even produce an additional displacement between the positive and negative charges.

A *nonpolar* molecule does not have this dipole arrangement until after a field is applied. The negative and positive charges shift in opposite directions against their mutual attraction and produce a dipole which is aligned with the electric field.

Either type of dipole may be described by its dipole moment **p**, as developed in Sec. 4.7, Eq. (37),

$$\mathbf{p} = Q\mathbf{d} \tag{18}$$

where Q is the positive one of the two bound charges composing the dipole, and **d** is the vector from the negative to the positive charge. We note again that the units of **p** are coulomb-meters.

If there are n dipoles per unit volume and we deal with a volume Δv, then there are $n\,\Delta v$ dipoles, and the total dipole moment is obtained by the vector sum,

$$\mathbf{p}_{\text{total}} = \sum_{i=1}^{n\,\Delta v} \mathbf{p}_i$$

If the dipoles are aligned in the same general direction, $\mathbf{p}_{\text{total}}$ may have a significant value. However, a random orientation may cause $\mathbf{p}_{\text{total}}$ to be essentially zero.

We now define the polarization **P** as the *dipole moment per unit volume,*

$$\mathbf{P} = \lim_{\Delta v \to 0} \frac{1}{\Delta v} \sum_{i=1}^{n\,\Delta v} \mathbf{p}_i \tag{19}$$

with units of coulombs per square meter. We shall treat **P** as a typical continuous field, even though it is obvious that it is essentially undefined at points within an atom or molecule. Instead, we should think of its value at any point as an average value taken over a sample volume Δv—large enough to contain many molecules ($n\,\Delta v$ in number), but yet sufficiently small to be considered incremental in concept.

Our immediate goal is to show that the bound volume charge density acts like the free volume charge density in producing an external field; we shall obtain a result similar to Gauss's law.

To be specific, let us assume that we have a dielectric containing nonpolar molecules. No molecule has a dipole moment, and $\mathbf{P} = 0$ throughout the material. Somewhere in the interior of the dielectric we select an incremental surface element $\Delta\mathbf{S}$, as shown in Fig. 5.9*a*, and apply an electric field **E**. The electric field produces a moment $\mathbf{p} = Q\mathbf{d}$ in each molecule, such that **p** and **d** make an angle θ with $\Delta\mathbf{S}$, as indicated in Fig. 5.9*b*.

Now let us inspect the movement of bound charges across $\Delta\mathbf{S}$. Each of the charges associated with the creation of a dipole must have moved a distance $\frac{1}{2}d\cos\theta$ in the direction perpendicular to $\Delta\mathbf{S}$. Thus, any positive charges initially lying below the surface $\Delta\mathbf{S}$ and within the distance $\frac{1}{2}d\cos\theta$ of the surface must have crossed $\Delta\mathbf{S}$ going upward. Also, any negative charges initially lying above

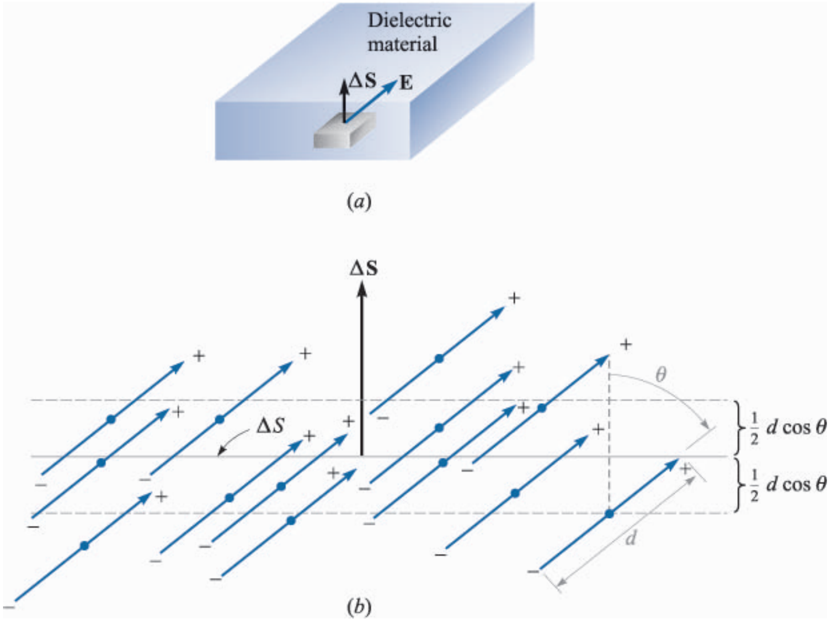


FIGURE 5.9
(a) An incremental surface element ΔS is shown in the interior of a dielectric in which an electric field $\mathbf{E}$ is present. (b) The nonpolar molecules form dipole moments $\mathbf{p}$ and a polarization $\mathbf{P}$. There is a net transfer of bound charge across $\Delta\mathbf{S}$.

the surface and within that distance $(\frac{1}{2}d\cos\theta)$ from $\Delta\mathbf{S}$ must have crossed $\Delta\mathbf{S}$ going downward. Therefore, since there are n molecules/m^3, the net total charge which crosses the elemental surface in an upward direction is equal to $nQd\cos\theta\Delta S$, or

$$\Delta Q_b = nQ\mathbf{d}\cdot\Delta\mathbf{S}$$

where the subscript on Q_b reminds us that we are dealing with a bound charge and not a free charge. In terms of the polarization, we have

$$\Delta Q_b = \mathbf{P}\cdot\Delta\mathbf{S}$$

If we interpret $\Delta\mathbf{S}$ as an element of a *closed* surface inside the dielectric material, then the direction of $\Delta\mathbf{S}$ is outward, and the net increase in the bound charge *within* the closed surface is obtained through the integral

$$Q_b = -\oint_S \mathbf{P}\cdot d\mathbf{S} \tag{20}$$

This last relationship has some resemblance to Gauss's law, and we may now generalize our definition of electric flux density so that it applies to media other

than free space. We first write Gauss's law in terms of $\epsilon_0\mathbf{E}$ and Q_T, the *total* enclosed charge, bound plus free:

$$Q_T = \oint_S \epsilon_0 \mathbf{E} \cdot d\mathbf{S} \tag{21}$$

where

$$Q_T = Q_b + Q$$

and Q is the total *free* charge enclosed by the surface S. Note that the free charge appears without subscript since it is the most important type of charge and will appear in Maxwell's equations.

Combining these last three equations, we obtain an expression for the free charge enclosed,

$$Q = Q_T - Q_b = \oint_S (\epsilon_0 \mathbf{E} + \mathbf{P}) \cdot d\mathbf{S} \tag{22}$$

We may now define $\mathbf{D}$ in more general terms than we did in Chap. 3,

$$\boxed{\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}} \tag{23}$$

There is thus an added term to D which appears when polarizable material is present. Thus,

$$Q = \oint_S \mathbf{D} \cdot d\mathbf{S} \tag{24}$$

where Q is the free charge enclosed.

Utilizing the several volume charge densities, we have

$$Q_b = \int_v \rho_b \, dv$$

$$Q = \int_v \rho_v \, dv$$

$$Q_T = \int_v \rho_T \, dv$$

With the help of the divergence theorem, we may therefore transform (20), (21), and (24) into the equivalent divergence relationships,

$$\nabla \cdot \mathbf{P} = -\rho_b$$

$$\nabla \cdot \epsilon_0 \mathbf{E} = \rho_T$$

$$\boxed{\nabla \cdot \mathbf{D} = \rho_v} \tag{25}$$

We shall emphasize only (24) and (25), the two expressions involving the free charge, in the work that follows.

In order to make any real use of these new concepts, it is necessary to know the relationship between the electric field intensity **E** and the polarization **P** which results. This relationship will, of course, be a function of the type of material, and we shall essentially limit our discussion to those isotropic materials for which **E** and **P** are linearly related. In an isotropic material the vectors **E** and **P** are always parallel, regardless of the orientation of the field. Although most engineering dielectrics are linear for moderate-to-large field strengths and are also isotropic, single crystals may be anisotropic. The periodic nature of crystalline materials causes dipole moments to be formed most easily along the crystal axes, and not necessarily in the direction of the applied field.

In *ferroelectric* materials the relationship between **P** and **E** is not only nonlinear, but also shows hysteresis effects; that is, the polarization produced by a given electric field intensity depends on the past history of the sample. Important examples of this type of dielectric are barium titanate, often used in ceramic capacitors, and Rochelle salt.

The linear relationship between **P** and **E** is

$$\boxed{\mathbf{P} = \chi_e \epsilon_0 \mathbf{E}} \tag{26}$$

where χ_e (chi) is a dimensionless quantity called the *electric susceptibility* of the material.

Using this relationship in (23), we have

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \chi_e \epsilon_0 \mathbf{E} = (\chi_e + 1)\epsilon_0 \mathbf{E}$$

The expression within the parentheses is now defined as

$$\epsilon_R = \chi_e + 1 \tag{27}$$

This is another dimensionless quantity and it is known as the *relative permittivity*, or *dielectric constant* of the material. Thus,

$$\mathbf{D} = \epsilon_0 \epsilon_R \mathbf{E} = \epsilon \mathbf{E} \tag{28}$$

where

$$\boxed{\epsilon = \epsilon_0 \epsilon_R} \tag{29}$$

and ϵ is the *permittivity*. The dielectric constants are given for some representative materials in Appendix C.

Anisotropic dielectric materials cannot be described in terms of a simple susceptibility or permittivity parameter. Instead, we find that each component of **D** may be a function of every component of **E**, and $\mathbf{D} = \epsilon \mathbf{E}$ becomes a matrix equation where **D** and **E** are each 3×1 column matrices and ϵ is a 3×3 square matrix. Expanding the matrix equation gives

$$D_x = \epsilon_{xx}E_x + \epsilon_{xy}E_y + \epsilon_{xz}E_z$$
$$D_y = \epsilon_{yx}E_x + \epsilon_{yy}E_y + \epsilon_{yz}E_z$$
$$D_z = \epsilon_{zx}E_x + \epsilon_{zy}E_y + \epsilon_{zz}E_z$$

Note that the elements of the matrix depend on the selection of the coordinate axes in the anisotropic material. Certain choices of axis directions lead to simpler matrices.[7]

D and **E** (and **P**) are no longer parallel, and although $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ remains a valid equation for anisotropic materials, we may continue to use $\mathbf{D} = \epsilon\mathbf{E}$ only by interpreting it as a matrix equation. We shall concentrate our attention on linear isotropic materials and reserve the general case for a more advanced text.

In summary, then, we now have a relationship between **D** and **E** which depends on the dielectric material present,

$$\boxed{\mathbf{D} = \epsilon\mathbf{E}} \tag{28}$$

where

$$\boxed{\epsilon = \epsilon_0\epsilon_R} \tag{29}$$

This electric flux density is still related to the free charge by either the point or integral form of Gauss's law:

$$\boxed{\nabla \cdot \mathbf{D} = \rho_v} \tag{25}$$

$$\boxed{\oint_S \mathbf{D} \cdot d\mathbf{S} = Q} \tag{24}$$

The use of the relative permittivity, as indicated by (29) above, makes consideration of the polarization, dipole moments, and bound charge unnecessary. However, when anisotropic or nonlinear materials must be considered, the relative permittivity, in the simple scalar form that we have discussed, is no longer applicable.

Let us now illustrate these new concepts with a numerical example.

---

[7] A more complete discussion of this matrix may be found in the Ramo, Whinnery, and Van Duzer reference listed at the end of this chapter.

Example 5.4

We locate a slab of Teflon in the region $0 \le x \le a$, and assume free space where $x < 0$ and $x > a$. Outside the Teflon there is a uniform field $\mathbf{E}_{out} = E_0 \mathbf{a}_x$ V/m. We seek values for **D**, **E**, and **P** everywhere.

Solution. The dielectric constant of Teflon is 2.1, and thus the electric susceptibility is 1.1.

Outside the slab, we have immediately $\mathbf{D}_{out} = \epsilon_0 E_0 \mathbf{a}_x$. Also since there is no dielectric material there, $\mathbf{P}_{out} = 0$. Now, any of the last four or five equations will enable us to relate the several fields inside the material to each other. Thus

$$\mathbf{D}_{in} = 2.1\epsilon_0 \mathbf{E}_{in} \qquad (0 \le x \le a)$$
$$\mathbf{P}_{in} = 1.1\epsilon_0 \mathbf{E}_{in} \qquad (0 \le x \le a)$$

As soon as we establish a value for any of these three fields within the dielectric, the other two can be found immediately. The difficulty lies in crossing over the boundary from the known fields external to the dielectric to the unknown ones within it. To do this we need a boundary condition, and this is the subject of the next exciting section. We shall complete this example then.

In the remainder of this text we shall describe polarizable materials in terms of **D** and ϵ rather than **P** and χ_e. We shall limit our discussion to isotropic materials.

✔ **D5.8.** A slab of dielectric material has a relative dielectric constant of 3.8 and contains a uniform electric flux density of 8 nC/m^2. If the material is lossless; find: (*a*) E; (*b*) P; (*c*) the average number of dipoles per cubic meter if the average dipole moment is 10^{-29} C · m

Ans. 238 V/m; 5.89 nC/m^2; 5.89×10^{20} m^{-3}

5.8 BOUNDARY CONDITIONS FOR PERFECT DIELECTRIC MATERIALS

How do we attack a problem in which there are two different dielectrics, or a dielectric and a conductor? This is another example of a *boundary condition*, such as the condition at the surface of a conductor whereby the tangential fields are zero and the normal electric flux density is equal to the surface charge density on the conductor. Now we take the first step in solving a two-dielectric problem, or a dielectric-conductor problem, by determining the behavior of the fields at the dielectric interface.

Let us first consider the interface between two dielectrics having permittivities ϵ_1 and ϵ_2 and occupying regions 1 and 2, as shown in Fig. 5.10. We first examine the tangential components by using

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0$$

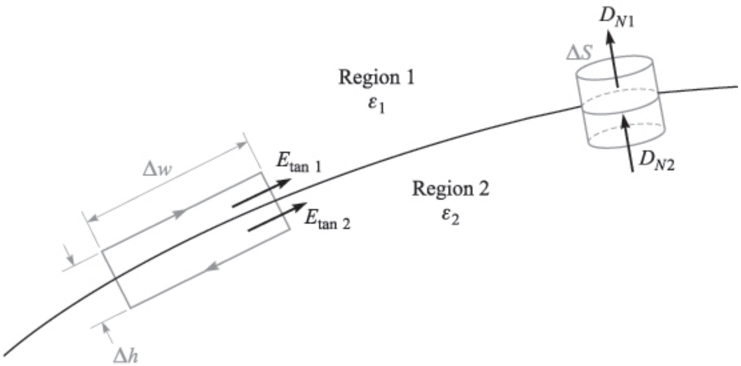


FIGURE 5.10
The boundary between perfect dielectrics of permittivities ϵ_1 and ϵ_2. The continuity of D_N is shown by the gaussian surface on the right, and the continuity of $E_{\tan}$ by the line integral about the closed path at the left.

around the small closed path on the left, obtaining

$$E_{\tan 1}\,\Delta w - E_{\tan 2}\,\Delta w = 0$$

The small contribution to the line integral by the normal component of **E** along the sections of length Δh becomes negligible as Δh decreases and the closed path crowds the surface. Immediately, then,

$$\boxed{E_{\tan 1} = E_{\tan 2}} \tag{30}$$

and we might feel that Kirchhoff's voltage law is still applicable to this case. Certainly we have shown that the potential difference between any two points on the boundary that are separated by a distance Δw is the same immediately above or below the boundary.

If the tangential electric field intensity is continuous across the boundary, then tangential **D** is discontinuous, for

$$\frac{D_{\tan 1}}{\epsilon_1} = E_{\tan 1} = E_{\tan 2} = \frac{D_{\tan 2}}{\epsilon_2}$$

or

$$\frac{D_{\tan 1}}{D_{\tan 2}} = \frac{\epsilon_1}{\epsilon_2} \tag{31}$$

The boundary conditions on the normal components are found by applying Gauss's law to the small "pillbox" shown at the right in Fig. 5.10. The sides are again very short, and the flux leaving the top and bottom surfaces is the difference

$$D_{N1}\Delta S - D_{N2}\Delta S = \Delta Q = \rho_S \Delta S$$

from which

$$\boxed{D_{N1} - D_{N2} = \rho_S} \tag{32}$$

What is this surface charge density? It cannot be a *bound* surface charge density, because we are taking the polarization of the dielectric into effect by using a dielectric constant different from unity; that is, instead of considering bound charges in free space, we are using an increased permittivity. Also, it is extremely unlikely that any *free* charge is on the interface, for no free charge is available in the perfect dielectrics we are considering. This charge must then have been placed there deliberately, thus unbalancing the total charge in and on this dielectric body. Except for this special case, then, we may assume ρ_S is zero on the interface and

$$\boxed{D_{N1} = D_{N2}} \tag{33}$$

or the normal component of D is continuous. It follows that

$$\epsilon_1 E_{N1} = \epsilon_2 E_{N2} \tag{34}$$

and normal **E** is discontinuous.

These conditions may be combined to show the change in the vectors **D** and **E** at the surface. Let $\mathbf{D}_1$ (and $\mathbf{E}_1$) make an angle θ_1 with a normal to the surface (Fig. 5.11). Since the normal components of **D** are continuous,

$$D_{N1} = D_1 \cos\theta_1 = D_2 \cos\theta_2 = D_{N2} \tag{35}$$

The ratio of the tangential components is given by (31) as

$$\frac{D_{\tan 1}}{D_{\tan 2}} = \frac{D_1 \sin\theta_1}{D_2 \sin\theta_2} = \frac{\epsilon_1}{\epsilon_2}$$

or

$$\epsilon_2 D_1 \sin\theta_1 = \epsilon_1 D_2 \sin\theta_2 \tag{36}$$

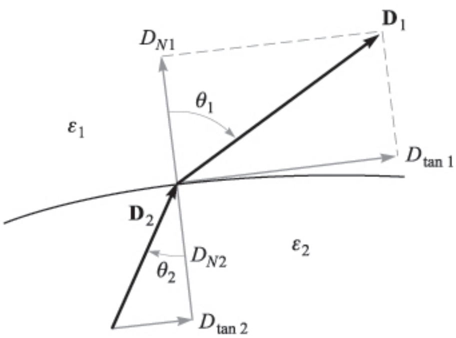


FIGURE 5.11
The refraction of **D** at a dielectric interface. For the case shown, $\epsilon_1 > \epsilon_2$; $\mathbf{E}_1$ and $\mathbf{E}_2$ are directed along $\mathbf{D}_1$ and $\mathbf{D}_2$, with $D_1 > D_2$ and $E_1 < E_2$.

and the division of this equation by (35) gives

$$\frac{\tan\theta_1}{\tan\theta_2} = \frac{\epsilon_1}{\epsilon_2} \tag{37}$$

In Fig. 5.11 we have assumed that $\epsilon_1 > \epsilon_2$, and therefore $\theta_1 > \theta_2$.

The direction of **E** on each side of the boundary is identical with the direction of **D**, because $\mathbf{D} = e\mathbf{E}$.

The magnitude of **D** in region 2 may be found from (35) and (36),

$$D_2 = D_1\sqrt{\cos^2\theta_1 + \left(\frac{\epsilon_2}{\epsilon_1}\right)^2 \sin^2\theta_1} \tag{38}$$

and the magnitude of $\mathbf{E}_2$ is

$$E_2 = E_1\sqrt{\sin^2\theta_1 + \left(\frac{\epsilon_1}{\epsilon_2}\right)^2 \cos^2\theta_1} \tag{39}$$

An inspection of these equations shows that D is larger in the region of larger permittivity (unless $\theta_1 = \theta_2 = 0^\circ$ where the magnitude is unchanged) and that E is larger in the region of smaller permittivity (unless $\theta_1 = \theta_2 = 90^\circ$, where its magnitude is unchanged).

These boundary conditions, (30), (31), (33), and (34), or the magnitude and direction relations derived from them, (37) to (39), allow us to find quickly the field on one side of a boundary *if we know the field on the other side*. In the example we began at the end of the previous section, this was the case. Now let's finish up that problem.

Example 5.5

Complete Example 5.4 by finding the fields within the Teflon ($\epsilon_R = 2.1$), given the uniform external field $\mathbf{E}_{out} = E_0\mathbf{a}_x$ in free space.

Solution. We recall that we had a slab of Teflon extending from $x = 0$ to $x = a$, as shown in Fig. 5.12, with free space on both sides of it and an external field $\mathbf{E}_{out} = E_0\mathbf{a}_x$. We also have $\mathbf{D}_{out} = \epsilon_0 E_0\mathbf{a}_x$ and $\mathbf{P}_{out} = 0$.

Inside, the continuity of D_N at the boundary allows us to find that $\mathbf{D}_{in} = \mathbf{D}_{out} = \epsilon_0 E_0\mathbf{a}_x$. This gives us $\mathbf{E}_{in} = \mathbf{D}_{in}/\epsilon = \epsilon_0 E_0\mathbf{a}_x/(\epsilon_R\epsilon_0) = 0.476E_0\mathbf{a}_x$. To get the polarization field in the dielectric, we use $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ and obtain

$$\mathbf{P}_{in} = \mathbf{D}_{in} - \epsilon_0\mathbf{E}_{in} = \epsilon_0 E_0\mathbf{a}_x - 0.476\epsilon_0 E_0\mathbf{a}_x = 0.524\epsilon_0 E_0\mathbf{a}_x$$

Summarizing then gives

$$\mathbf{D}_{in} = \epsilon_0 E_0\mathbf{a}_x \qquad (0 \le x \le a)$$
$$\mathbf{E}_{in} = 0.476E_0\mathbf{a}_x \qquad (0 \le x \le a)$$
$$\mathbf{P}_{in} = 0.524\epsilon_0 E_0\mathbf{a}_x \qquad (0 \le x \le a)$$

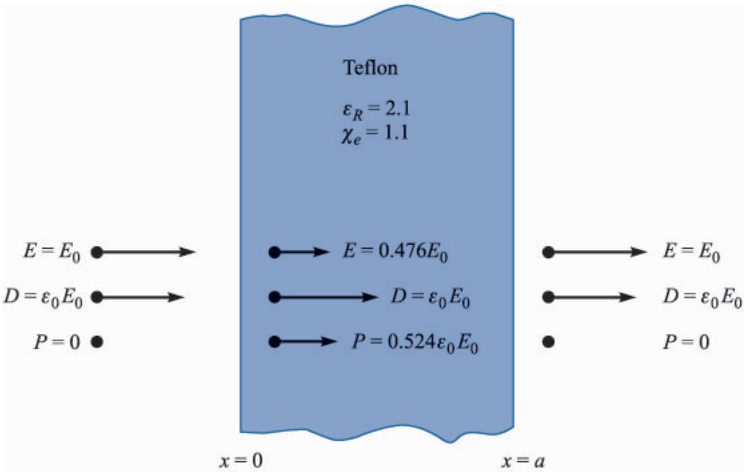


FIGURE 5.12
A knowledge of the electric field external to the dielectric enables us to find the remaining external fields first, and then to use the continuity of normal **D** to begin finding the internal fields.

A practical problem most often does not provide us with a direct knowledge of the field on either side of the boundary. The boundary conditions must be used to help us determine the fields on both sides of the boundary from the other information which is given. A simple problem of this type will be considered in Sec. 5.10.

The boundary conditions existing at the interface between a conductor and a dielectric are much simpler than those above. First, we know that **D** and **E** are both zero inside the conductor. Second, the tangential **E** and **D** field components must both be zero to satisfy

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0$$

and

$$\mathbf{D} = \epsilon \mathbf{E}$$

Finally, the application of Gauss's law,

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = Q$$

shows once more that both **D** and **E** are normal to the conductor surface and that $D_N = \rho_S$ and $E_N = \rho_S/\epsilon$. We see, then, that the boundary conditions we developed previously for the conductor-free space boundary are valid for the conductor-dielectric boundary if we replace ϵ_0 by ϵ. Thus

$$\boxed{D_t = E_t = 0} \tag{40}$$

$$\boxed{D_N = \epsilon E_N = \rho_S} \tag{41}$$

It is interesting to spend a moment discovering how any change that is introduced *internally* within a conducting material arrives at the surface as a surface charge. We should understand that this is not a common occurrence, but it does give us some additional insight into the characteristics of a conductor,

Given Ohm's law,

$$\mathbf{J} = \sigma \mathbf{E}$$

and the continuity equation,

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho_v}{\partial t}$$

in which $\mathbf{J}$ and ρ_v both involve only free charges, we have

$$\nabla \cdot \sigma \mathbf{E} = -\frac{\partial \rho_v}{\partial t}$$

or

$$\nabla \cdot \frac{\sigma}{\epsilon} \mathbf{D} = -\frac{\partial \rho_v}{\partial t}$$

If we assume that the medium is homogeneous, so that σ and ϵ are not functions of position,

$$\nabla \cdot \mathbf{D} = -\frac{\epsilon}{\sigma} \frac{\partial \rho_v}{\partial t}$$

Now we may use Maxwell's first equation to obtain

$$\rho_v = -\frac{\epsilon}{\sigma} \frac{\partial \rho_v}{\partial t}$$

Let us now make the simplifying assumption that σ is not a function of ρ_v. This is probably not a very good assumption, for we found in Sec. 5.3, Eq. (9), that σ depended on both ρ_v and the mobility, but it leads to an easy solution that at least permits us to compare different conductors. We simply rearrange and integrate directly, obtaining

$$\rho_v = \rho_0 e^{-(\sigma/\epsilon)t}$$

where ρ_0 = charge density at $t = 0$. This shows an exponential decay of charge density at every point with a time constant of ϵ/σ. This time constant, often called the *relaxation time*, may be calculated for a relatively poor conductor, such as distilled water, from the data in Appendix C, giving

$$\frac{\epsilon}{\sigma} = \frac{80 \times 8.854 \times 10^{-12}}{2 \times 10^{-4}} = 3.54 \quad \mu\text{s}$$

In 3.54 μs any charge we place in the interior of a body of distilled water has dropped to about 37 percent of its initial value. This rapid decay is characteristic of good conductors and we see that, except for an extremely short transient period, we may safely consider the charge density to be zero within a good conductor.

With the physical materials with which we must work, no dielectric material is without some few free electrons; all have conductivities different from zero, and charge introduced internally in any of them will eventually reach the surface.

With the knowledge we now have of conducting materials, dielectric materials, and the necessary boundary conditions, we are ready to define and discuss capacitance.

✔ **D5.9.** Let the region $z < 0$ be composed of a uniform dielectric material for which $\epsilon_R = 3.2$, while the region $z > 0$ is characterized by $\epsilon_R = 2$. Let $\mathbf{D}_1 = -30\mathbf{a}_x + 50\mathbf{a}_y + 70\mathbf{a}_z$ nC/m² and find: (*a*) D_{N1}; (*b*) $\mathbf{D}_{t1}$; (*c*) D_{t1}; (*d*) D_1; (*e*) θ_1; (*f*) $\mathbf{P}_1$.

Ans. 70 nC/m²; $-30\mathbf{a}_x + 50\mathbf{a}_y$ nC/m²; 58.3 nC/m²; 91.1 nC/m²; 39.8°; $-20.6\mathbf{a}_x + 34.4\mathbf{a}_y + 48.1\mathbf{a}_z$ nC/m²

✔ **D5.10.** Continue Prob. D5.9 by finding: (*a*) $\mathbf{D}_{N2}$; (*b*) $\mathbf{D}_{t2}$; (*c*) $\mathbf{D}_2$; (*d*) $\mathbf{P}_2$; (*e*) θ_2

Ans. $70\mathbf{a}_z$ nC/m²; $-18.75\mathbf{a}_x + 31.25\mathbf{a}_y$ nC/m²; $-18.75\mathbf{a}_x + 31.25\mathbf{a}_y + 70\mathbf{a}_z$ nC/m²; $-9.38\mathbf{a}_x + 15.63\mathbf{a}_y + 35\mathbf{a}_z$ nC/m²; 27.5°

5.9 CAPACITANCE

Now let us consider two conductors embedded in a homogeneous dielectric (Fig. 5.13). Conductor M_2 carries a total positive charge Q, and M_1 carries an equal

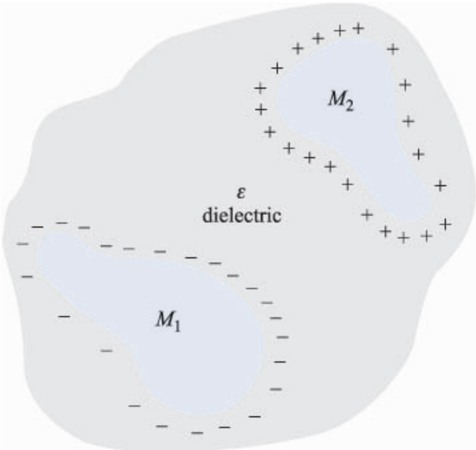


FIGURE 5.13
Two oppositely charged conductors M_1 and M_2 surrounded by a uniform dielectric. The ratio of the magnitude of the charge on either conductor to the magnitude of the potential difference between them is the capacitance C.

negative charge. There are no other charges present, and the *total* charge of the system is zero.

We now know that the charge is carried on the surface as a surface charge density and also that the electric field is normal to the conductor surface. Each conductor is, moreover, an equipotential surface. Since M_2 carries the positive charge, the electric flux is directed from M_2 to M_1, and M_2 is at the more positive potential. In other words, work must be done to carry a positive charge from M_1 to M_2.

Let us designate the potential difference between M_2 and M_1 as V_0. We may now define the *capacitance* of this two-conductor system as the ratio of the magnitude of the total charge on either conductor to the magnitude of the potential difference between conductors,

$$\boxed{C = \frac{Q}{V_0}} \tag{42}$$

In general terms, we determine Q by a surface integral over the positive conductors, and we find V_0 by carrying a unit positive charge from the negative to the positive surface,

$$C = \frac{\oint_S \epsilon \mathbf{E} \cdot d\mathbf{S}}{-\int_-^+ \mathbf{E} \cdot d\mathbf{L}} \tag{43}$$

The capacitance is independent of the potential and total charge, for their ratio is constant. If the charge density is increased by a factor of N, Gauss's law indicates that the electric flux density or electric field intensity also increases by N, as does the potential difference. The capacitance is a function only of the physical dimensions of the system of conductors and of the permittivity of the homogeneous dielectric.

Capacitance is measured in *farads* (F), where a farad is defined as one coulomb per volt. Common values of capacitance are apt to be very small fractions of a farad, and consequently more practical units are the microfarad (μF), the nanofarad (nF), and the picofarad (pF).

We can apply the definition of capacitance to a simple two-conductor system in which the conductors are identical, infinite parallel planes with separation d (Fig. 5.14). Choosing the lower conducting plane at $z = 0$ and the upper one at $z = d$, a uniform sheet of surface charge $\pm\rho_S$ on each conductor leads to the uniform field [Sec. 2.5, Eq. (22)]

$$\mathbf{E} = \frac{\rho_S}{\epsilon}\mathbf{a}_z$$

where the permittivity of the homogeneous dielectric is ϵ, and

$$\mathbf{D} = \rho_S \mathbf{a}_z$$

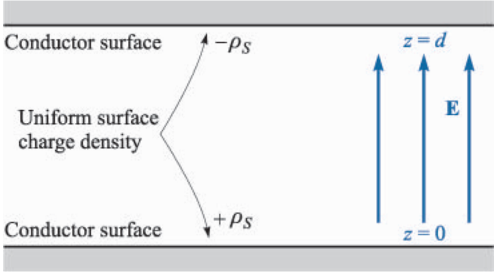


FIGURE 5.14
The problem of the parallel-plate capacitor. The capacitance per square meter of surface area is ϵ/d.

The charge on the lower plane must then be positive, since **D** is directed upward, and the normal value of **D**,

$$D_N = D_z = \rho_S$$

is equal to the surface charge density there. On the upper plane,

$$D_N = -D_z$$

and the surface charge there is the negative of that on the lower plane.

The potential difference between lower and upper planes is

$$V_0 = -\int_{\text{upper}}^{\text{lower}} \mathbf{E} \cdot d\mathbf{L} = -\int_d^0 \frac{\rho_S}{\epsilon} dz = \frac{\rho_S}{\epsilon} d$$

Since the total charge on either plane is infinite, the capacitance is infinite. A more practical answer is obtained by considering planes, each of area S, whose linear dimensions are much greater than their separation d. The electric field and charge distribution are then almost uniform at all points not adjacent to the edges, and this latter region contributes only a small percentage of the total capacitance, allowing us to write the familiar result

$$Q = \rho_S S$$

$$V_0 = \frac{\rho_S}{\epsilon} d$$

$$\boxed{C = \frac{Q}{V_0} = \frac{\epsilon S}{d}} \tag{44}$$

More rigorously, we might consider (44) as the capacitance of a portion of the infinite-plane arrangement having a surface area S. Methods of calculating the effect of the unknown and nonuniform distribution near the edges must wait until we are able to solve more complicated potential problems.

Example 5.6

Calculate the capacitance of a parallel-plate capacitor having a mica dielectric, $\epsilon_R = 6$, a plate area of $10\,\text{in}^2$, and a separation of 0.01 in.

Solution. We may find that

$$S = 10 \times 0.0254^2 = 6.45 \times 10^{-3} \quad \text{m}^2$$
$$d = 0.01 \times 0.0254 = 2.54 \times 10^{-4} \quad \text{m}$$

and therefore

$$C = \frac{6 \times 8.854 \times 10^{-12} \times 6.45 \times 10^{-3}}{2.54 \times 10^{-4}} = 1.349 \quad \text{nF}$$

A large plate area is obtained in capacitors of small physical dimensions by stacking smaller plates in 50- or 100-decker sandwiches, or by rolling up foil plates separated by a flexible dielectric.

Table *C*.1 in Appendix *C* also indicates that materials are available having dielectric constants greater than 1000.

If more than two conductors are involved, *partial capacitances* between each pair of conductors must be defined. This is interestingly discussed in Maxwell's works.[8]

Finally, the total energy stored in the capacitor is

$$W_E = \tfrac{1}{2}\int_{\text{vol}} \epsilon E^2\, dv = \tfrac{1}{2}\int_0^S \int_0^d \frac{\epsilon \rho_S^2}{\epsilon^2}\, dz\, dS = \tfrac{1}{2}\frac{\rho_S^2}{\epsilon} Sd = \tfrac{1}{2}\frac{\epsilon S}{d}\frac{\rho_S^2 d^2}{\epsilon^2}$$

or

$$\boxed{W_E = \tfrac{1}{2} C V_0^2 = \tfrac{1}{2} Q V_0 = \tfrac{1}{2}\frac{Q^2}{C}} \qquad (45)$$

which are all familiar expressions. Equation (45) also indicates that the energy stored in a capacitor with a fixed potential difference across it increases as the dielectric constant of the medium increases.

D5.11. Find the relative permittivity of the dielectric material present in a parallel-plate capacitor if: (*a*) $S = 0.12\,\text{m}^2$, $d = 80\,\mu\text{m}$, $V_0 = 12\,\text{V}$, and the capacitor contains $1\,\mu\text{J}$ of energy; (*b*) the stored energy density is $100\,\text{J/m}^3$, $V_0 = 200\,\text{V}$, and $d = 45\,\mu\text{m}$; (*c*) $E = 200\,\text{kV/m}$, $\rho_S = 20\,\mu\text{C/m}^2$, and $d = 100\,\mu\text{m}$.

Ans. 6.27; 1.144; 11.29

[8] See the Suggested References at the end of the chapter.

5.10 SEVERAL CAPACITANCE EXAMPLES

As a first brief example we choose a coaxial cable or coaxial capacitor of inner radius a, outer radius b, and length L. No great derivational struggle is required, because the potential difference is given as Eq. (11) in Sec. 4.3, and we find the capacitance very simply by dividing this by the total charge $\rho_L L$ in the length L. Thus,

$$C = \frac{2\pi\epsilon L}{\ln(b/a)} \tag{46}$$

Next we consider a spherical capacitor formed of two concentric spherical conducting shells of radius a and b, $b > a$. The expression for the electric field was obtained previously by Gauss's law,

$$E_r = \frac{Q}{4\pi\epsilon r^2}$$

where the region between the spheres is a dielectric with permittivity ϵ. The expression for potential difference was found from this by the line integral [Sec. 4.3, Eq. (12)]. Thus,

$$V_{ab} = \frac{Q}{4\pi\epsilon}\left(\frac{1}{a} - \frac{1}{b}\right)$$

Here Q represents the total charge on the inner sphere, and the capacitance becomes

$$C = \frac{Q}{V_{ab}} = \frac{4\pi\epsilon}{\dfrac{1}{a} - \dfrac{1}{b}} \tag{47}$$

If we allow the outer sphere to become infinitely large, we obtain the capacitance of an isolated spherical conductor,

$$C = 4\pi\epsilon a \tag{48}$$

For a diameter of 1 cm, or a sphere about the size of a marble,

$$C = 0.556 \quad \text{pF}$$

in free space.

Coating this sphere with a different dielectric layer, for which $\epsilon = \epsilon_1$, extending from $r = a$ to $r = r_1$,

$$D_r = \frac{Q}{4\pi r^2}$$

$$E_r = \frac{Q}{4\pi\epsilon_1 r^2} \qquad (a < r < r_1)$$

$$= \frac{Q}{4\pi\epsilon_0 r^2} \qquad (r_1 < r)$$

and the potential difference is

$$V_a - V_\infty = -\int_{r_1}^{a} \frac{Q\,dr}{4\pi\epsilon_1 r^2} - \int_{\infty}^{r_1} \frac{Q\,dr}{4\pi\epsilon_0 r^2}$$

$$= \frac{Q}{4\pi}\left[\frac{1}{\epsilon_1}\left(\frac{1}{a} - \frac{1}{r_1}\right) + \frac{1}{\epsilon_0 r_1}\right]$$

Therefore,

$$C = \frac{4\pi}{\dfrac{1}{\epsilon_1}\left(\dfrac{1}{a} - \dfrac{1}{r_1}\right) + \dfrac{1}{\epsilon_0 r_1}} \tag{49}$$

In order to look at the problem of multiple dielectrics a little more thoroughly, let us consider a parallel-plate capacitor of area S and spacing d, with the usual assumption that d is small compared to the linear dimensions of the plates. The capacitance is $\epsilon_1 S/d$, using a dielectric of permittivity ϵ_1. Now let us replace a part of this dielectric by another of permittivity ϵ_2, placing the boundary between the two dielectrics parallel to the plates (Fig. 5.15).

Some of us may immediately suspect that this combination is effectively two capacitors in series, yielding a total capacitance of

$$C = \frac{1}{\dfrac{1}{C_1} + \dfrac{1}{C_2}}$$

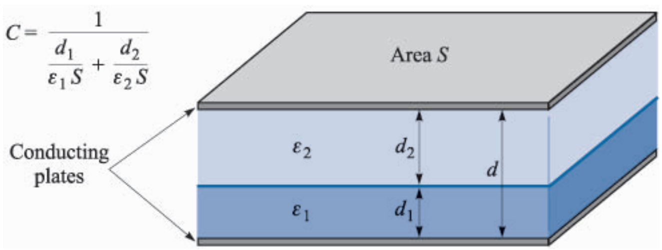


FIGURE 5.15
A parallel-plate capacitor containing two dielectrics with the dielectric interface parallel to the conducting plates.

where $C_1 = \epsilon_1 S/d_1$ and $C_2 = \epsilon_2 S/d_2$. This is the correct result, but we can obtain it using less intuition and a more basic approach.

Since our capacitance definition, $C = Q/V$, involves a charge and a voltage, we may assume either and then find the other in terms of it. The capacitance is not a function of either, but only of the dielectrics and the geometry. Suppose we assume a potential difference V_0 between the plates. The electric field intensities in the two regions, E_2 and E_1, are both uniform, and $V_0 = E_1 d_1 + E_2 d_2$. At the dielectric interface, E is normal and $D_{N1} = D_{N2}$, or $\epsilon_1 E_1 = \epsilon_2 E_2$. Eliminating E_2 in our V_0 relation, we have

$$E_1 = \frac{V_0}{d_1 + d_2(\epsilon_1/\epsilon_2)}$$

and the surface charge density therefore has the magnitude

$$\rho_{S1} = D_1 = \epsilon_1 E_1 = \frac{V_0}{\dfrac{d_1}{\epsilon_1} + \dfrac{d_2}{\epsilon_2}}$$

Since $D_1 = D_2$, the magnitude of the surface charge is the same on each plate. The capacitance is then

$$C = \frac{Q}{V_0} = \frac{\rho_S S}{V_0} = \frac{1}{\dfrac{d_1}{\epsilon_1 S} + \dfrac{d_2}{\epsilon_2 S}} = \frac{1}{\dfrac{1}{C_1} + \dfrac{1}{C_2}}$$

As an alternate (and slightly simpler) solution, we might assume a charge Q on one plate, leading to a charge density Q/S and a value of D that is also Q/S. This is true in both regions, as $D_{N1} = D_{N2}$ and D is normal. Then $E_1 = D/\epsilon_1 = Q/(\epsilon_1 S)$, $E_2 = D/\epsilon_2 = Q/(\epsilon_2 S)$, and the potential differences across the regions are $V_1 = E_1 d_1 = Q d_1/(\epsilon_1 S)$, and $V_2 = E_2 d_2 = Q d_2/(\epsilon_2 S)$. The capacitance is

$$C = \frac{Q}{V} = \frac{Q}{V_1 + V_2} = \frac{1}{\dfrac{d_1}{\epsilon_1 S} + \dfrac{d_2}{\epsilon_2 S}} \tag{50}$$

How would the method of solution or the answer change if there were a third conducting plane along the interface? We would now expect to find surface charge on each side of this conductor, and the magnitudes of these charges should be equal. In other words, we think of the electric lines not as passing directly from one outer plate to the other, but as terminating on one side of this interior plane and then continuing on the other side. The capacitance is unchanged, provided, of course, that the added conductor is of negligible thickness. The addition of a thick conducting plate will increase the capacitance if the separation of the outer plates is kept constant, and this is an example of a more general theorem which states that the replacement of any portion of the dielectric by a conducting body will cause an increase in the capacitance.

If the dielectric boundary were placed *normal* to the two conducting plates and the dielectrics occupied areas of S_1 and S_2, then an assumed potential difference V_0 would produce field strengths $E_1 = E_2 = V_0/d$. These are tangential fields at the interface, and they must be equal. Then we may find in succession D_1, D_2, ρ_{S1}, ρ_{S2}, and Q, obtaining a capacitance

$$C = \frac{\epsilon_1 S_1 + \epsilon_2 S_2}{d} = C_1 + C_2 \tag{51}$$

as we should expect.

At this time we can do very little with a capacitor in which two dielectrics are used in such a way that the interface is not everywhere normal or parallel to the fields. Certainly we know the boundary conditions at each conductor and at the dielectric interface; however, we do not know the fields to which to apply the boundary conditions. Such a problem must be put aside until our knowledge of field theory has increased and we are willing and able to use more advanced mathematical techniques.

✔ **D5.12.** Determine the capacitance of: (*a*) a 1-ft length of 35B/U coaxial cable, which has an inner conductor 0.1045 in in diameter, a polyethylene dielectric ($\epsilon_R = 2.26$ from Table C.1), and an outer conductor which has an inner diameter of 0.680 in; (*b*) a conducting sphere of radius 2.5 mm, covered with a polyethylene layer 2 mm thick, surrounded by a conducting sphere of radius 4.5 mm; (*c*) two rectangular conducting plates, 1 cm by 4 cm, with negligible thickness, between which are three sheets of dielectric, each 1 cm by 4 cm, and 0.1 mm thick, having dielectric constants of 1.5, 2.5, and 6.

Ans. 20.5 pF; 1.414 pF; 28.0 pF

5.11 CAPACITANCE OF A TWO-WIRE LINE

We conclude this chapter with the problem of the two-wire line. The final configuration will consist of two parallel conducting cylinders, each of circular cross section, and we shall be able to find complete information about the electric field intensity, the potential field, the surface-charge-density distribution, and the capacitance. This arrangement is an important type of transmission line, as is the coaxial cable we have discussed several times before.

We begin by investigating the potential field of two infinite line charges. Fig 5.16 shows a positive line charge in the xz plane at $x = a$ and a negative line charge at $x = -a$. The potential of a single line charge with zero reference at a radius of R_0 is

$$V = \frac{\rho_L}{2\pi\epsilon} \ln \frac{R_0}{R}$$

We now write the expression for the combined potential field in terms of the radial distances from the positive and negative lines, R_1 and R_2, respectively,

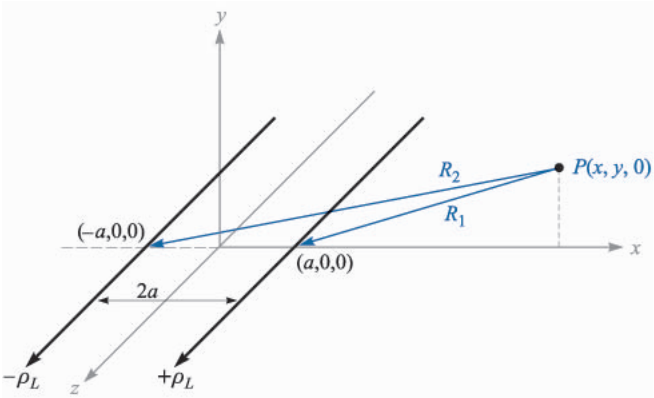


FIGURE 5.16
Two parallel infinite line charges carrying opposite charge. The positive line is at $x = a$, $y = 0$, and the negative line is at $x = -a$, $y = 0$. A general point $P(x, y, 0)$ in the xy plane is radially distant R_1 and R_2 from the positive and negative lines, respectively. The equipotential surfaces are circular cylinders.

$$V = \frac{\rho_L}{2\pi\epsilon}\left(\ln\frac{R_{10}}{R_1} - \ln\frac{R_{20}}{R_2}\right) = \frac{\rho_L}{2\pi\epsilon}\ln\frac{R_{10}R_2}{R_{20}R_1}$$

We choose $R_{10} = R_{20}$, thus placing the zero reference at equal distances from each line. This surface is the $x = 0$ plane. Expressing R_1 and R_2 in terms of x and y,

$$V = \frac{\rho_L}{2\pi\epsilon}\ln\sqrt{\frac{(x+a)^2 + y^2}{(x-a)^2 + y^2}} = \frac{\rho_L}{4\pi\epsilon}\ln\frac{(x+a)^2 + y^2}{(x-a)^2 + y^2} \tag{52}$$

In order to recognize the equipotential surfaces and adequately understand the problem we are going to solve, some algebraic manipulations are necessary. Choosing an equipotential surface $V = V_1$, we define K_1 as a dimensionless parameter that is a function of the potential V_1,

$$K_1 = e^{4\pi\epsilon V_1/\rho_L} \tag{53}$$

so that

$$K_1 = \frac{(x+a)^2 + y^2}{(x-a)^2 + y^2}$$

After multiplying and collecting like powers, we obtain

$$x^2 - 2ax\frac{K_1 + 1}{K_1 - 1} + y^2 + a^2 = 0$$

We next work through a couple of lines of algebra and complete the square,

$$\left(x - a\frac{K_1 + 1}{K_1 - 1}\right)^2 + y^2 = \left(\frac{2a\sqrt{K_1}}{K_1 - 1}\right)^2$$

This shows that the $V = V_1$ equipotential surface is independent of z (or is a cylinder) and intersects the xy plane in a circle of radius b,

$$b = \frac{2a\sqrt{K_1}}{K_1 - 1}$$

which is centered at $x = h$, $y = 0$, where

$$h = a\frac{K_1 + 1}{K_1 - 1}$$

Now let us attack a physical problem by considering a zero-potential conducting plane located at $x = 0$, and a conducting cylinder of radius b and potential V_0 with its axis located a distance h from the plane. We solve the last two equations for a and K_1 in terms of the dimensions b and h,

$$a = \sqrt{h^2 - b^2} \tag{54}$$

and

$$\sqrt{K_1} = \frac{h + \sqrt{h^2 - b^2}}{b} \tag{55}$$

But the potential of the cylinder is V_0, so (53) leads to

$$\sqrt{K_1} = e^{2\pi\epsilon V_0/\rho_L}$$

Therefore,

$$\rho_L = \frac{4\pi\epsilon V_0}{\ln K_1} \tag{56}$$

Thus, given h, b, and V_0, we may determine a, ρ_L, and the parameter K_1. The capacitance between the cylinder and plane is now available. For a length L in the z direction, we have

$$C = \frac{\rho_L L}{V_0} = \frac{4\pi\epsilon L}{\ln K_1} = \frac{2\pi\epsilon L}{\ln\sqrt{K_1}}$$

or

$$C = \frac{2\pi\epsilon L}{\ln[h + \sqrt{h^2 - b^2}/b]} = \frac{2\pi\epsilon L}{\cosh^{-1}(h/b)} \tag{57}$$

The heavy black circle in Fig. 5.17 shows the cross section of a cylinder of 5-m radius at a potential of 100 V in free space, with its axis 13 m distant from a plane at zero potential. Thus, $b = 5$, $h = 13$, $V_0 = 100$, and we rapidly find the location of the equivalent line charge from (54),

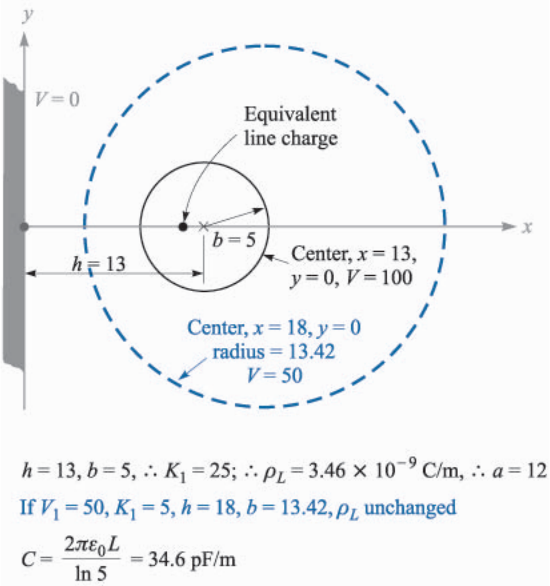


FIGURE 5.17
A numerical example of the capacitance, linear charge density, position of an equivalent line charge, and characteristics of the mid-equipotential surface for a cylindrical conductor of 5-m radius at a potential of 100 V, parallel to and 13 m from a conducting plane at zero potential.

$$a = \sqrt{h^2 - b^2} = \sqrt{13^2 - 5^2} = 12\,\text{m}$$

the value of the potential parameter K_1 from (55),

$$\sqrt{K_1} = \frac{h + \sqrt{h^2 - b^2}}{b} = \frac{13 + 12}{5} = 5 \qquad K_1 = 25$$

the strength of the equivalent line charge from (56)

$$\rho_L = \frac{4\pi\epsilon V_0}{\ln K_1} = \frac{4\pi \times 8.854 \times 10^{-12} \times 100}{\ln 25} = 3.46 \quad \text{nC/m}$$

and the capacitance between cylinder and plane from (57),

$$C = \frac{2\pi\epsilon}{\cosh^{-1}(h/b)} = \frac{2\pi \times 8.854 \times 10^{-12}}{\cosh^{-1}(13/5)} = 34.6 \quad \text{pF/m}$$

We may also identify the cylinder representing the 50-V equipotential surface by finding new values for K_1, h, and b. We first use (53) to obtain

$$K_1 = e^{4\pi\epsilon V_1/\rho_L} = e^{4\pi \times 8.854 \times 10^{-12} \times 50/3.46 \times 10^{-9}} = 5.00$$

Then the new radius is

$$b = \frac{2a\sqrt{K_1}}{K_1 - 1} = \frac{2 \times 12\sqrt{5}}{5 - 1} = 13.42 \quad \text{m}$$

and the corresponding value of h becomes

$$h = a\frac{K_1 + 1}{K_1 - 1} = 12\frac{5+1}{5-1} = 18 \quad \text{m}$$

This cylinder is shown in color in Fig. 5.17.

The electric field intensity can be found by taking the gradient of the potential field, as given by (52),

$$\mathbf{E} = -\nabla\left[\frac{\rho_L}{4\pi\epsilon}\ln\frac{(x+a)^2 + y^2}{(x-a)^2 + y^2}\right]$$

Thus,

$$\mathbf{E} = -\frac{\rho_L}{4\pi\epsilon}\left[\frac{2(x+a)\mathbf{a}_x + 2y\mathbf{a}_y}{(x+a)^2 + y^2} - \frac{2(x-a)\mathbf{a}_x + 2y\mathbf{a}_y}{(x-a)^2 + y^2}\right)$$

and

$$\mathbf{D} = e\mathbf{E} = -\frac{\rho_L}{2\pi}\left[\frac{(x+a)\mathbf{a}_x + y\mathbf{a}_y}{(x+a)^2 + y^2} - \frac{(x-a)\mathbf{a}_x + y\mathbf{a}_y}{(x-a)^2 + y^2}\right]$$

If we evaluate D_x at $x = h - b$, $y = 0$, we may obtain $\rho_{S,max}$

$$\rho_{S,max} = -D_{x,x=h-b,y=0} = \frac{\rho_L}{2\pi}\left[\frac{h-b+a}{(h-b+a)^2} - \frac{h-b-a}{(h-b-a)^2}\right]$$

For our example,

$$\rho_{S,max} = \frac{3.46 \times 10^{-9}}{2\pi}\left[\frac{13-5+12}{(13-5+12)^2} - \frac{13-5-12}{(13-5-12)^2}\right] = 0.1650 \quad \text{nC/m}^2$$

Similarly, $\rho_{S,min} = D_{x,x=h+b,y=0}$, and

$$\rho_{S,min} = \frac{3.46 \times 10^{-9}}{2\pi}\left[\frac{13+5+12}{30^2} - \frac{13+5-12}{6^2}\right] = 0.0734 \quad \text{nC/m}^2$$

Thus,

$$\rho_{S,max} = 2.25\rho_{S,min}$$

If we apply (57) to the case of a conductor for which $b \ll h$, then

$$\ln[(h + \sqrt{h^2 - b^2}/b] \doteq \ln([h + h)/b] \doteq \ln(2h/b)$$

and

$$C = \frac{2\pi\epsilon L}{\ln(2h/b)} \quad (b \ll h) \tag{58}$$

The capacitance between two circular conductors separated by a distance $2h$ is one-half the capacitance given by (57) or (58). This last answer is of interest

because it gives us an expression for the capacitance of a section of two-wire transmission line, one of the types of transmission lines studied later in Chap. 12.

✔ **D5.13.** A conducting cylinder with a radius of 1 cm and at a potential of 20 V is parallel to a conducting plane which is at zero potential. The plane is 5 cm distant from the cylinder axis. If the conductors are embedded in a perfect dielectric for which $\epsilon_R = 4.5$, find: (*a*) the capacitance per unit length between cylinder and plane; (*b*) $\rho_{S,\max}$ on the cylinder.

Ans. 109.2 pF/m; 2.21 nC/m

SUGGESTED REFERENCES

1. Adler, R. B., A. C. Smith, and R. L. Longini: "Introduction to Semiconductor Physics," John Wiley & Sons, Inc., New York, 1964. Semiconductor theory is treated at an undergraduate level.
2. Dekker, A. J.: "Electrical Engineering Materials," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1959. This admirable little book covers dielectrics, conductors, semiconductors, and magnetic materials.
3. Fano, R. M., L. J. Chu, and R. B. Adler: "Electromagnetic Fields, Energy, and Forces," John Wiley & Sons, Inc., New York, 1960. Polarization in dielectrics is discussed in the first part of chap. 5. This junior-level text presupposes a full-term physics course in electricity and magnetism, and it is therefore slightly more advanced in level. The introduction beginning on p. 1 should be read.
4. Fink, D. G., and H. W. Beaty: "Standard Handbook for Electrical Engineers," 12th ed., McGraw-Hill Book Company, New York, 1987.
5. Matsch, L. W.: "Capacitors, Magnetic Circuits, and Transformers," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964. Many of the practical aspects of capacitors are discussed in chap. 2.
6. Maxwell, J. C.: "A Treatise on Electricity and Magnetism," 3d ed., Oxford University Press, New York, 1904, or an inexpensive paperback edition, Dover Publications, Inc., New York, 1954.
7. Ramo, S., J. R. Whinnery, and T. Van Duzer: "Fields and Waves in Communications Electronics," 3rd ed., John Wiley & Sons, Inc., New York, 1994. This book is essentially the fifth edition of the senior authors' popular texts of 1944 and 1953. Although it is directed primarily toward beginning graduate students, it may be profitably read by anyone who is familiar with basic electromagnetic concepts. Anisotropic dielectric materials are discussed on pp. 699–712.
8. Wert, C. A., and R. M. Thomson: "Physics of Solids," 2d ed., McGraw-Hill Book Company, New York, 1970. This is an advanced undergraduate-level text that covers metals, semiconductors, and dielectrics.

PROBLEMS

5.1 Given the current density $\mathbf{J} = -10^4(\sin 2x\, e^{-2y}\mathbf{a}_x + \cos 2x\, e^{-2y}\mathbf{a}_y)$ kA/m^2: (*a*) find the total current crossing the plane $y = 1$ in the $\mathbf{a}_y$ direction in the region $0 < x < 1$, $0 < z < 2$. Find the total current leaving the region $0 < x, y < 1$, $2 < z < 3$ by: (*b*) integrating $\mathbf{J} \cdot \mathbf{dS}$ over the surface of the cube; (*c*) employing the divergence theorem.

5.2 Let the current density be $\mathbf{J} = 2\rho \cos^2 \phi\, \mathbf{a}_\rho - \rho \sin 2\phi\, \mathbf{a}_\phi$ A/m^2 within the region $2.1 < \rho < 2.5$, $0 < \phi < 0.1$ rad, $6 < z < 6.1$. Find the total current I crossing the surface: (*a*) $\rho = 2.2$, $0 < \phi < 0.1$, $6 < z < 6.1$ in the $\mathbf{a}_\rho$ direction; (*b*) $\phi = 0.05$, $2.2 < \rho < 2.5$, $6 < z < 6.1$, in the $\mathbf{a}_\phi$ direction. (*c*) Evaluate $\nabla \cdot \mathbf{J}$ at $P(\rho = 2.4, \phi = 0.08, z = 6.05)$.

5.3 Let $\mathbf{J} = \dfrac{400 \sin \theta}{r^2 + 4}$ A/m^2. (*a*) Find the total current flowing through that portion of the spherical surface $r = 0.8$ bounded by $\theta = 0.1\pi$, $\theta = 0.3\pi$, $0 < \phi < 2\pi$. (*b*) Find the average value of J over the defined area.

5.4 The cathode of a planar vacuum tube is at $z = 0$. Let $\mathbf{E} = -4 \times 10^6 \mathbf{a}_z$ V/m for $z > 0$. An electron ($e = 1.602 \times 10^{-19}$ C, $m = 9.11 \times 10^{-31}$ kg) is emitted from the cathode with zero initial velocity at $t = 0$. (*a*) Find $v(t)$. (*b*) Find $z(t)$, the electron location as a function of time. (*c*) Determine $v(z)$. (*d*) Make the assumption that electrons are emitted continuously as a beam with a 0.25-mm radius and a total current of 60 µA. Find $\mathbf{J}(z)$ and $\rho_v(z)$.

5.5 Let $\mathbf{J} = \dfrac{25}{\rho}\mathbf{a}_\rho - \dfrac{20}{\rho^2 + 0.01}\mathbf{a}_z$ A/m^2, and: (*a*) find the total current crossing the plane $z = 0.2$ in the $\mathbf{a}_z$ direction for $\rho < 0.4$. (*b*) Calculate $\dfrac{\partial \rho_v}{\partial t}$. (*c*) Find the total outward current crossing the closed surface defined by $\rho = 0.01$, $\rho = 0.4$, $z = 0$, and $z = 0.2$. (*d*) Show that the divergence theorem is satisified for $\mathbf{J}$ and the surface specified.

5.6 Let $\epsilon = \epsilon_0$ and $V = 90z^{4/3}$ in the region $z = 0$. (*a*) Obtain expressions for $\mathbf{E}$, $\mathbf{D}$, and ρ_v as functions of z. (*b*) If the velocity of the charge density is given as $v_x = 5 \times 10^6 z^{2/3}$ m/s, find J_z at $z = 0$ and $z = 0.1$ m.

5.7 Assuming that there is no transformation of mass to energy or vice versa, it is possible to write a continuity equation for mass. (*a*) If we use the continuity equation for charge as our model, what quantities correspond to $\mathbf{J}$ and ρ_v? (*b*) Given a cube 1 cm on a side, experimental data show that the rates at which mass is leaving each of the six faces are 10.25, −9.85, 1.75, −2.00, −4.05, and 4.45 mg/s. If we assume that the cube is an incremental volume element, determine an approximate value for the time rate of change of density at its center.

5.8 The continuity equation for mass equates the divergence of the mass rate of flow (mass per second per square meter) to the negative of the density (mass per cubic meter). After setting up a cartesian coordinate system inside a star, Captain Kirk and his intrepid crew make measurements over the faces of a cube centered at the origin with edges 40 km long and

parallel to the coordinate axes. They find the mass rate of flow of material outward across the six faces to be −1112, 1183, 201, −196, 1989, and $-1920\,\mathrm{kg/km^2}\cdot\mathrm{s}$. (*a*) Estimate the divergence of the mass rate of flow at the origin. (*b*) Estimate the rate of change of the density at the origin.

5.9 (*a*) Using data tabulated in Appendix C, calculate the required diameter for a 2-m long nichrome wire that will dissipate an average power of 450 W when 120-V rms at 60 Hz is applied to it. (*b*) Calculate the rms current density in the wire.

5.10 A steel wire has a radius of 2 mm and a conductivity of 6×10^6 S/m. The steel wire has an aluminum ($\sigma = 3.8 \times 10^7$ S/m) coating of 2-mm thickness. Let the total current carried by this hybrid conductor be 80 A dc. Find: (*a*) J_{st}; (*b*) J_{A1}; (*c*) E_{st}; (*d*) E_{Al}; (*e*) the voltage between the ends of the conductor if it is 1 mi long.

5.11 Two perfectly conducting cylindrical surfaces are located at $\rho = 3$ and $\rho = 5$ cm. The total current passing radially outward through the medium between the cylinders is 3 A dc. (*a*) Find the voltage and resistance between the cylinders, and **E** in the region between the cylinders, if a conducting material having $\sigma = 0.05$ S/m is present for $3 < \rho < 5$ cm. (*b*) Show that integrating the power dissipated per unit volume over the volume gives the total dissipated power.

5.12 The spherical surfaces $r = 3$ and $r = 5$ cm are perfectly conducting, and the total current passing radially outward through the medium between the surfaces is 3 A dc. (*a*) Find the voltage and resistance between the spheres, and **E** in the region between them, if a conducting material having $\sigma = 0.05$ S/m is present for $3 < r < 5$ cm. (*b*) Repeat if $\sigma = 0.0005/r$ for $3 < r < 5$ cm. (*c*) Show that integrating the power dissipated per unit volume in part *b* over the volume gives the total dissipated power.

5.13 A hollow cylindrical tube with a rectangular cross section has external dimensions of 0.5 in by 1 in and a wall thickness of 0.05 in. Assume that the material is brass for which $\sigma = 1.5 \times 10^7$ S/m. A current of 200 A dc is flowing down the tube. (*a*) What voltage drop is present across a 1 m length of the tube? (*b*) Find the voltage drop if the interior of the tube is filled with a conducting material for which $\sigma = 1.5 \times 10^5$ S/m.

5.14 Find the magnitude of the electric field intensity in a conductor if: (*a*) the current density is $5\,\mathrm{MA/m^2}$, the electron mobility is $3 \times 10^{-3}\,\mathrm{m^2/V\cdot s}$, and the volume charge density is $-2.4 \times 10^{10}\,\mathrm{C/m^3}$; (*b*) $J = 3\,\mathrm{MA/m^2}$ and the resistivity is $3 \times 10^{-8}\,\Omega\cdot\mathrm{m}$.

5.15 Let $V = 10(\rho + 1)z^2 \cos\phi$ V in free space. (*a*) Let the equipotential surface $V = 20$ V define a conductor surface. Find the equation of the conductor surface. (*b*) Find ρ and **E** at that point on the conductor surface where $\phi = 0.2\pi$ and $z = 1.5$. (*c*) Find $|\rho_S|$ at that point.

5.16 A potential field in free space is given as $V = (80 \cos\theta \sin\phi)/r^3$ V. Point $P(r = 2, \theta = \pi/3, \phi = \pi/2)$ lies on a conducting surface. (*a*) Write the equation of the conducting surface. (*b*) Find a unit normal directed out-

ward to the surface, assuming the origin is inside the surface. (*c*) Find **E** at *P*.

5.17 Given the potential field $V = \dfrac{100xz}{x^2+4}$ V in free space: (*a*) find **D** at the surface $z = 0$. (*b*) Show that the $z = 0$ surface is an equipotential surface. (*c*) Assume that the $z = 0$ surface is a conductor and find the total charge on that portion of the conductor defined by $0 < x < 2$, $-3 < y < 0$.

5.18 Let us assume a field $\mathbf{E} = 3y^2z^3\mathbf{a}_x + 6xyz^3\mathbf{a}_y + 9xy^2z^2$ V/m in free space, and also assume that point $P(2, 1, 0)$ lies on a conducting surface. (*a*) Find ρ_v just adjacent to the surface at *P*. (*b*) Find ρ_S at *P*. (*c*) Show that $V = -3xy^2z^3$ V. (*d*) Determine V_{PQ}, given $Q(1, 1, 1)$.

5.19 Let $V = 20x^2yz - 10z^2$ V in free space. (*a*) Determine the equations of the equipotential surfaces on which $V = 0$ and 60 V. (*b*) Assume these are conducting surfaces and find the surface charge density at that point on the $V = 60$-V surface where $x = 2$ and $z = 1$. It is known that $0 \le V \le 60$ V is the field-containing region. (*c*) Give the unit vector at this point that is normal to the conducting surface and directed toward the $V = 0$ surface.

5.20 A conducting plane is located at $z = 0$ in free space, and a 20-nC point charge is present at $Q(2, 4, 6)$. (*a*) If $V = 0$ at $z = 0$, find *V* at $P(5, 3, 1)$. (*b*) Find **E** at *P*. (*c*) Find ρ_S at $A(5, 3, 0)$.

5.21 Let the surface $y = 0$ be a perfect conductor in free space. Two uniform infinite line charges of 30 nC/m each are located at $x = 0$, $y = 1$, and $x = 0$, $y = 2$. (*a*) Let $V = 0$ at the plane $y = 0$, and find *V* at $P(1, 2, 0)$. (*b*) Find **E** at *P*.

5.22 Let the plane $x = 0$ be a perfect conductor in free space. Locate a point charge of 4 nC at $P_1(7, 1, -2)$, and a point charge of -3 nC at $P_2(4, 2, 1)$. (*a*) Find **E** at $A(5, 0, 0)$. (*b*) Find $|\rho_S|$ at $B(3, 0, 0)$.

5.23 A dipole with $\mathbf{p} = 0.1\mathbf{a}_z\ \mu\text{C}\cdot\text{m}$ is located at $A(1, 0, 0)$ in free space, and the $x = 0$ plane is perfectly conducting. (*a*) Find *V* at $P(2, 0, 1)$. (*b*) Find the equation of the 200-V equipotential surface in cartesian coordinates.

5.24 The mobilities for intrinsic silicon at a certain temperature are $\mu_e = 0.14\ \text{m}^2/\text{V}\cdot\text{s}$ and $\mu_h = 0.035\ \text{m}^2/\text{V}\cdot\text{s}$. The concentration of both holes and electrons is $2.2 \times 10^{16}\ \text{m}^{-3}$. Determine both the conductivity and resistivity of this silicon sample.

5.25 Electron and hole concentrations increase with temperature. For pure silicon suitable expressions are $\rho_h = -\rho_e = 6200T^{1.5}e^{-7000/T}\ \text{C/m}^3$. The functional dependence of the mobilities on temperature is given by $\mu_h = 2.3 \times 10^5 T^{-2.7}\ \text{m}^2/\text{V}\cdot\text{s}$ and $\mu_e = 2.1 \times 10^5 T^{-2.5}\ \text{m}^2/\text{V}\cdot\text{s}$. Find σ at: (*a*) 0° C; (*b*) 40° C; (*c*) 80° C.

5.26 A little donor impurity, such as arsenic, is added to pure silicon so that the electron concentration is 2×10^{17} conduction electrons per cubic meter while the number of holes per cubic meter is only 1.1×10^{15}. If $\mu_e = 0.15\ \text{m}^2/\text{V}\cdot\text{s}$ for this sample, and $\mu_h = 0.045\ \text{m}^2/\text{V}\cdot\text{s}$, determine the conductivity and resistivity.

5.27 Atomic hydrogen contains 5.5×10^{25} atoms/m^3 at a certain temperature and pressure. When an electric field of 4 kV/m is applied, each dipole formed by the electron and the positive nucleus has an effective length of 7.1×10^{-19} m. Find: (*a*) P; (*b*) ϵ_R.

5.28 In a certain region where the relative permittivity is 2.4, $\mathbf{D} = 2\mathbf{a}_x - 4\mathbf{a}_y + 5\mathbf{a}_z$ nC/m^2. Find: (*a*) $\mathbf{E}$; (*b*) $\mathbf{P}$; (*c*) $|\nabla V|$.

5.29 A coaxial conductor has radii $a = 0.8$ mm and $b = 3$ mm and a polystyrene dielectric for which $\epsilon_R = 2.56$. If $\mathbf{P} = \frac{2}{\rho}\mathbf{a}_\rho$ nC/m^2 in the dielectric, find: (*a*) $\mathbf{D}$ and $\mathbf{E}$ as functions of ρ; (*b*) V_{ab} and χ_e. (*c*) If there are 4×10^{19} molecules per cubic meter in the dielectric, find $\mathbf{p}(\rho)$.

5.30 Given the potential field $V = 200 - 50x + 20y$ V in a dielectric material for which $\epsilon_R = 2.1$, find: (*a*) $\mathbf{E}$; (*b*) $\mathbf{D}$; (*c*) $\mathbf{P}$; (*d*) ρ_v; (*e*) ρ_b; (*f*) ρ_T.

5.31 The surface $x = 0$ separates two perfect dielectrics. For $x > 0$ let $\epsilon_R = \epsilon_{R1} = 3$, while $\epsilon_{R2} = 5$ where $x < 0$. If $\mathbf{E}_1 = 80\mathbf{a}_x - 60\mathbf{a}_y - 30\mathbf{a}_z$ V/m, find: (*a*) E_{N1}; (*b*) $\mathbf{E}_{t1}$; (*c*) $\mathbf{E}_{t1}$; (*d*) E_1; (*e*) the angle θ_1 between $\mathbf{E}_1$ and a normal to the surface; (*f*) D_{N2}; (*g*) D_{t2}; (*h*) $\mathbf{D}_2$; (*i*) $\mathbf{P}_2$ (*j*) the angle θ_2 between $\mathbf{E}_2$ and a normal to the surface.

5.32 In Fig. 5.18 let $\mathbf{D}_1 = 3\mathbf{a}_x - 4\mathbf{a}_y + 5\mathbf{a}_z$ nC/m^2 and find: (*a*) $\mathbf{D}_2$; (*b*) $\mathbf{D}_{N2}$; (*c*) $\mathbf{D}_{t2}$; (*d*) the energy density in each region; (*e*) the angle that $\mathbf{D}_2$ makes with $\mathbf{a}_z$; (*f*) D_2/D_1; (*g*) P_2/P_1.

5.33 Two perfect dielectrics have relative permittivities $\epsilon_{R1} = 2$ and $\epsilon_{R2} = 8$. The planar interface between them is the surface $x - y + 2z = 5$. The origin lies in region 1. If $\mathbf{E}_1 = 100\mathbf{a}_x + 200\mathbf{a}_y - 50\mathbf{a}_z$ V/m, find $\mathbf{E}_2$.

5.34 Let the spherical surfaces $r = 4$ cm and $r = 9$ cm be separated by two perfect dielectric shells, $\epsilon_{R1} = 2$ for $4 < r < 6$ cm, and $\epsilon_{R2} = 5$ for $6 < r < 9$ cm. If $\mathbf{E}_1 = \frac{2000}{r^2}\mathbf{a}_r$ V/m, find: (*a*) $\mathbf{E}_2$; (*b*) the total electrostatic energy stored in each region.

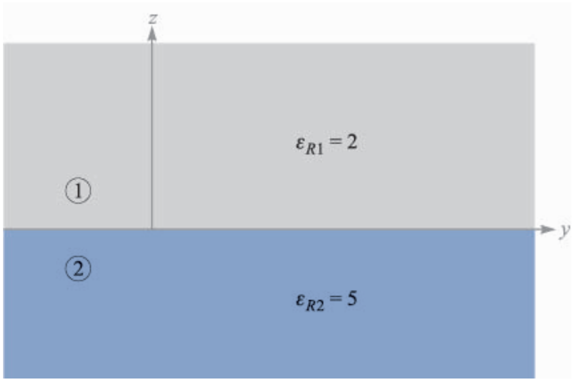


FIGURE 5.18
See Prob. 32.

5.35 Let the cylindrical surfaces $\rho = 4\,\text{cm}$ and $\rho = 9\,\text{cm}$ enclose two wedges of perfect dielectrics, $\epsilon_{R1} = 2$ for $0 < \phi < \pi/2$, and $\epsilon_{R2} = 5$ for $\pi/2 < \phi < 2\pi$. If $\mathbf{E}_1 = \dfrac{2000}{\rho}\mathbf{a}_\rho$ V/m, find: (*a*) $\mathbf{E}_2$; (*b*) the total electrostatic energy stored in a 1-m length of each region.

5.36 Let $S = 120\,\text{cm}^2$, $d = 4\,\text{mm}$, and $\epsilon_R = 12$ for a parallel-plate capacitor. (*a*) Calculate the capacitance. (*b*) After connecting a 40-V battery across the capacitor, calculate E, D, Q, and the total stored electrostatic energy. (*c*) The source is now removed and the dielectric carefully withdrawn from between the plates. Again calculate E, D, Q, and the energy. (*d*) What is the voltage between the plates?

5.37 Capacitors tend to be more expensive as their capacitance and maximum voltage V_{max} increase. The voltage V_{max} is limited by the field strength at which the dielectric breaks down, E_{BD}. Which of these dielectrics will give the largest CV_{max} product for equal plate areas: (*a*) air: $\epsilon_R = 1$, $E_{BD} = 3\,\text{MV/m}$; (*b*) barium titanate: $\epsilon_R = 1200$, $E_{BD} = 3\,\text{MV/m}$; (*c*) silicon dioxide: $\epsilon_R = 3.78$, $E_{BD} = 16\,\text{MV/m}$; (*d*) polyethylene: $\epsilon_R = 2.26$, $E_{BD} = 4.7\,\text{MV/m}$.

5.38 A dielectric circular cylinder used between the plates of a capacitor has a thickness of 0.2 mm and a radius of 1.4 cm. The dielectric properties are $\epsilon_R = 400$ and $\sigma = 10^{-5}$ S/m. (*a*) Calculate C. (*b*) Find the quality factor $Q_{QF}(Q_{QF} = \omega RC)$ of the capacitor at $f = 10\,\text{kHz}$. (*c*) If the maximum field strength permitted in the dielectric is 2 MV/m, what is the maximum permissible voltage across the capacitor? (*d*) What energy is stored when this voltage is applied?

5.39 A parallel-plate capacitor is filled with a nonuniform dielectric characterized by $\epsilon_R = 2 + 2 \times 10^6 x^2$, where x is the distance from one plate. If $S = 0.02\,\text{m}^2$ and $d = 1\,\text{mm}$, find C.

5.40 (*a*) The width of the region containing ϵ_{R1} in Fig. 5.19 is 1.2 m. Find ϵ_{R1} if $\epsilon_{R2} = 2.5$ and the total capacitance is 60 nF. (*b*) Find the width of each region (containing ϵ_{R1} and ϵ_{R2}) if $C_{\text{total}} = 80\,\text{nF}$, $\epsilon_{R2} = 3\epsilon_{R1}$, and $C_1 = 2C_2$.

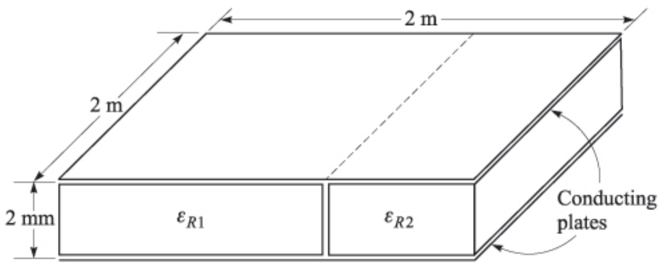


FIGURE 5.19
See Prob. 40.

5.41 Let $\epsilon_{R1} = 2.5$ for $0 < y < 1$ mm, $\epsilon_{R2} = 4$ for $1 < y < 3$ mm, and ϵ_{R3} for $3 < y < 5$ mm. Conducting surfaces are present at $y = 0$ and $y = 5$ mm. Calculate the capacitance per square meter of surface area if: (*a*) ϵ_{R3} is air; (*b*) $\epsilon_{R3} = \epsilon_{R1}$; (*c*) $\epsilon_{R3} = \epsilon_{R2}$; (*d*) ϵ_{R3} is silver.

5.42 Cylindrical conducting surfaces are located at $\rho = 0.8$ cm and 3.6 cm. The region 0.8 cm $< \rho < a$ contains a dielectric for which $\epsilon_R = 4$, while $\epsilon_R = 2$ for $a < \rho < 3.6$ cm. (*a*) Find a so that the voltage across each dielectric layer is the same. (*b*) Find the total capacitance per meter.

5.43 Two coaxial conducting cylinders of radius 2 cm and 4 cm have a length of 1 m. The region between the cylinders contains a layer of dielectric from $\rho = c$ to $\rho = d$ with $\epsilon_R = 4$. Find the capacitance if: (*a*) $c = 2$ cm, $d = 3$ cm; (*b*) $d = 4$ cm, and the volume of dielectric is the same as in part *a*.

5.44 Conducting cylinders lie at $\rho = 3$ and 12 mm; both extend from $z = 0$ to $z = 1$ m. Perfect dielectrics occupy the interior region: $\epsilon_R = 1$ for $3 < \rho < 6$ mm, $\epsilon_R = 4$ for $6 < \rho < 9$ mm, and $\epsilon_R = 8$ for $9 < \rho < 12$ mm. (*a*) Calculate C. (*b*) If the voltage between the cylinders is 100 V, plot $|E_\rho|$ versus ρ.

5.45 Two conducting spherical shells have radii $a = 3$ cm and $b = 6$ cm. The interior is a perfect dielectric for which $\epsilon_R = 8$. (*a*) Find C. (*b*) A portion of the dielectric is now removed so that $\epsilon_R = 1, 0 < \phi < \pi/2$, and $\epsilon_R = 8$, $\pi/2 < \phi < 2\pi$. Again find C.

5.46 Conducting cylinders lie at $\rho = 3$ and 12 mm; both extend from $z = 0$ to $z = 1$ m. Perfect dielectrics occupy the interior region: $\epsilon_R = 1$ for $3 < \rho < 6$ mm, $\epsilon_R = 4$ for $6 < \rho < 9$ mm, and $\epsilon_R = 8$ for $9 < \rho < 12$ mm. (*a*) Calculate C. (*b*) If the voltage between the cylinders is 100 V, plot $|E_\rho|$ versus ρ.

5.47 With reference to Fig. 5.17, let $b = 6$ m, $h = 15$ m, and the conductor potential be 250 V. Take $\epsilon = \epsilon_0$. Find values for K_1, ρ_L, a, and C.

5.48 A potential function in free space is given by $V = -20 + 10 \ln \dfrac{(5+y)^2 + x^2}{(5-y)^2 + x^2}$ V. Describe: (*a*) the 0-V equipotential surface; (*b*) the 10-V equipotential surface.

5.49 A 2-cm-diameter conductor is suspended in air with its axis 5 cm from a conducting plane. Let the potential of the cylinder be 100 V and that of the plane be 0 V. Find the surface charge density on the: (*a*) cylinder at a point nearest the plane; (*b*) plane at a point nearest the cylinder.

CHAPTER

6

EXPERIMENTAL MAPPING METHODS

We have seen in the last few chapters that the potential is the gateway to any information we desire about the electrostatic field at a point. The path is straight, and travel on it is easy in whichever direction we wish to go. The electric field intensity may be found from the potential by the gradient operation, which is a differentiation, and the electric field intensity may then be used to find the electric flux density by multiplying by the permittivity. The divergence of the flux density, again a differentiation, gives the volume charge density; and the surface charge density on any conductors in the field is quickly found by evaluating the flux density at the surface. Our boundary conditions show that it must be normal to such a surface.

Integration is still required if we need more information than the value of a field or charge density *at a point*. Finding the total charge on a conductor, the total energy stored in an electrostatic field, or a capacitance or resistance value are examples of such problems, each requiring an integration. These integrations cannot generally be avoided, no matter how extensive our knowledge of field theory becomes, and indeed, we should find that the greater this knowledge becomes, the more integrals we should wish to evaluate. Potential can do one important thing for us, and that is to quickly and easily furnish us with the quantity we must integrate.

Our goal, then, is to find the potential first. This cannot be done in terms of a charge configuration in a practical problem, because no one is kind enough to

tell us exactly how the charges are distributed. Instead, we are usually given several conducting objects or conducting boundaries and the potential difference between them. Unless we happen to recognize the boundary surfaces as belonging to a simple problem we have already disposed of, we can do little now and must wait until Laplace's equation is discussed in the following chapter.

Although we thus postpone the mathematical solution to this important type of practical problem, we may acquaint ourselves with several experimental methods of finding the potential field. Some of these methods involve special equipment such as an electrolytic trough, a fluid-flow device, resistance paper and the associated bridge equipment, or rubber sheets; others use only pencil, paper, and a good supply of erasers. The *exact* potential can never be determined, but sufficient accuracy for engineering purposes can usually be attained. One other method, called the *iteration* method, does allow us to achieve any desired accuracy for the potential, but the number of calculations required increases very rapidly as the desired accuracy increases.

Several of the experimental methods to be described below are based on an analogy with the electrostatic field, rather than directly on measurements on this field itself.

Finally, we cannot introduce this subject of experimental methods of finding potential fields without emphasizing the fact that many practical problems possess such a complicated geometry that no exact method of finding that field is possible or feasible and experimental techniques are the only ones which can be used.

6.1 CURVILINEAR SQUARES

Our first mapping method is a graphical one, requiring only pencil and paper. Besides being economical, it is also capable of yielding good accuracy if used skillfully and patiently. Fair accuracy (5 to 10 percent on a capacitance determination) may be obtained by a beginner who does no more than follow the few rules and hints of the art.

The method to be described is applicable only to fields in which no variation exists in the direction normal to the plane of the sketch. The procedure is based on several facts we have already demonstrated:

1. A conductor boundary is an equipotential surface.
2. The electric field intensity and electric flux density are both perpendicular to the equipotential surfaces.
3. **E** and **D** are therefore perpendicular to the conductor boundaries and possess zero tangential values.
4. The lines of electric flux, or streamlines, begin and terminate on charge and hence, in a charge-free, homogeneous dielectric, begin and terminate only on the conductor boundaries.

Let us consider the implications of these statements by drawing the streamlines on a sketch which already shows the equipotential surfaces. In Fig. 6.1*a* two conductor boundaries are shown, and equipotentials are drawn with a constant potential difference between lines. We should remember that these lines are only the cross sections of the equipotential surfaces, which are cylinders (although not circular), since no variation in the direction normal to the surface of the paper is permitted. We arbitrarily choose to begin a streamline, or flux line, at A on the surface of the more positive conductor. It leaves the surface normally and must cross at right angles the undrawn but very real equipotential surfaces between the conductor and the first surface shown. The line is continued to the other conductor, obeying the single rule that the intersection with each equipotential must be square. Turning the paper from side to side as the line progresses enables us to maintain perpendicularity more accurately. The line has been completed in Fig. 6.1*b*.

In a similar manner, we may start at B and sketch another streamline ending at B'. Before continuing, let us interpret the meaning of this pair of streamlines. The streamline, by definition, is everywhere tangent to the electric field intensity or to the electric flux density. Since the streamline is tangent to the electric flux density, the flux density is tangent to the streamline and no electric flux may cross any streamline. In other words, if there is a charge of 5 μC on the surface between A and B (and extending 1 m into the paper), then 5 μC of flux begins in this region and all must terminate between A' and B'. Such a pair of lines is sometimes called a flux *tube*, because it physically seems to carry flux from one conductor to another without losing any.

We now wish to construct a third streamline, and both the mathematical and visual interpretations we may make from the sketch will be greatly simplified if we draw this line starting from some point C chosen so that the same amount of flux is carried in the tube BC as is contained in AB. How do we choose the position of C?

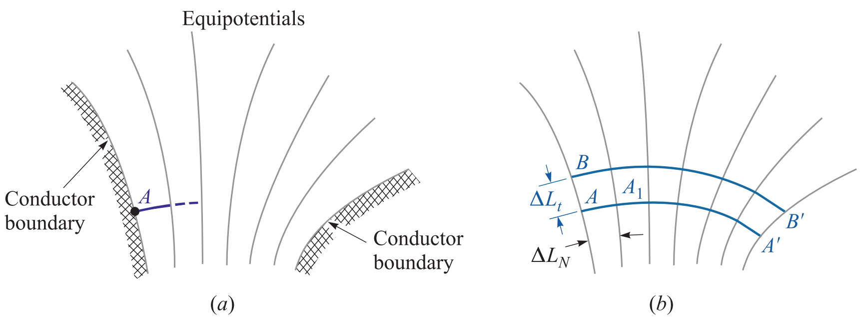


FIGURE 6.1
(*a*) Sketch of the equipotential surfaces between two conductors. The increment of potential between each of the two adjacent equipotentials is the same. (*b*) One flux line has been drawn from A to A', and a second from B to B'.

The electric field intensity at the midpoint of the line joining A to B may be found approximately by assuming a value for the flux in the tube AB, say $\Delta\Psi$, which allows us to express the electric flux density by $\Delta\Psi/\Delta L_t$, where the depth of the tube into the paper is 1 m and ΔL_t is the length of the line joining A to B. The magnitude of E is then

$$E = \frac{1}{\epsilon}\frac{\Delta\Psi}{\Delta L_t}$$

However, we may also find the magnitude of the electric field intensity by dividing the potential difference between points A and A_1, lying on two adjacent equipotential surfaces, by the distance from A to A_1. If this distance is designated ΔL_N and an increment of potential between equipotentials of ΔV is assumed, then

$$E = \frac{\Delta V}{\Delta L_N}$$

This value applies most accurately to the point at the middle of the line segment from A to A_1, while the previous value was most accurate at the midpoint of the line segment from A to B. If, however, the equipotentials are close together (ΔV small) and the two streamlines are close together ($\Delta\Psi$ small), the two values found for the electric field intensity must be approximately equal,

$$\frac{1}{\epsilon}\frac{\Delta\Psi}{\Delta L_t} = \frac{\Delta V}{\Delta L_N} \tag{1}$$

Throughout our sketch we have assumed a homogeneous medium (ϵ constant), a constant increment of potential between equipotentials (ΔV constant), and a constant amount of flux per tube ($\Delta\Psi$ constant). To satisfy all these conditions, (1) shows that

$$\frac{\Delta L_t}{\Delta L_N} = \text{constant} = \frac{1}{\epsilon}\frac{\Delta\Psi}{\Delta V} \tag{2}$$

A similar argument might be made at any point in our sketch, and we are therefore led to the conclusion that a constant ratio must be maintained between the distance between streamlines as measured along an equipotential, and the distance between equipotentials as measured along a streamline. It is this *ratio* which must have the same value at every point, not the individual lengths. Each length must decrease in regions of greater field strength, because ΔV is constant.

The simplest ratio we can use is unity, and the streamline from B to B' shown in Fig. 6.1*b* was started at a point for which $\Delta L_t = \Delta L_N$. Since the ratio of these distances is kept at unity, the streamlines and equipotentials divide the field-containing region into curvilinear squares, a term implying a planar geometric figure which differs from a true square in having slightly curved and slightly unequal sides but which approaches a square as its dimensions decrease.

Those incremental surface elements in our three coordinate systems which are planar may also be drawn as curvilinear squares.

We may now rapidly sketch in the remainder of the streamlines by keeping each small box as square as possible. The complete sketch is shown in Fig. 6.2.

The only difference between this example and the production of a field map using the method of curvilinear squares is that the intermediate potential surfaces are not given. The streamlines and equipotentials must both be drawn on an original sketch which shows only the conductor boundaries. Only one solution is possible, as we shall prove later by the uniqueness theorem for Laplace's equation, and the rules we have outlined above are sufficient. One streamline is begun, an equipotential line is roughed in, another streamline is added, forming a curvilinear square, and the map is gradually extended throughout the desired region. Since none of us can ever expect to be perfect at this, we shall soon find that we can no longer make squares and also maintain right-angle corners. An error is accumulating in the drawing, and our present troubles should indicate the nature of the correction to make on some of the earlier work. It is usually best to start again on a fresh drawing, with the old one available as a guide.

The construction of a useful field map is an art; the science merely furnishes the rules. Proficiency in any art requires practice. A good problem for beginners is the coaxial cable or coaxial capacitor, since all the equipotentials are circles, while the flux lines are straight lines. The next sketch attempted should be two parallel circular conductors, where the equipotentials are again circles, but with different centers. Each of these is given as a problem at the end of the chapter, and the accuracy of the sketch may be checked by a capacitance calculation as outlined below.

Fig. 6.3 shows a completed map for a cable containing a square inner conductor surrounded by a circular conductor. The capacitance is found from $C = Q/V_0$ by replacing Q by $N_Q \Delta Q = N_Q \Delta \Psi$, where N_Q is the number of flux tubes joining the two conductors, and letting $V_0 = N_V \Delta V$, where N_V is the number of potential increments between conductors,

$$C = \frac{N_Q \Delta Q}{N_V \Delta V}$$

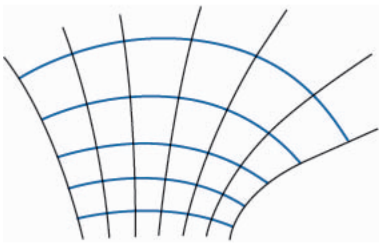


FIGURE 6.2
The remainder of the streamlines have been added to Fig. 6.1*b* by beginning each new line normally to the conductor and maintaining curvilinear squares throughout the sketch.

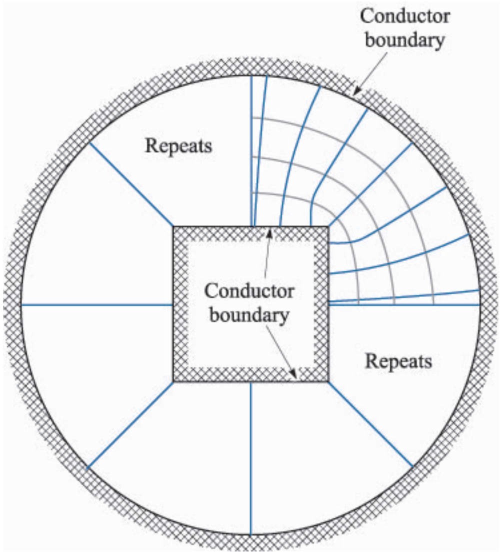


FIGURE 6.3
An example of a curvilinear-square field map. The side of the square is two thirds the radius of the circle. $N_V = 4$ and $N_Q = 8 \times 3.25 \times 26$, and therefore $C = \epsilon_0 N_Q / N_V = 57.6\,\text{pF/m}$.

and then using (2),

$$C = \frac{N_Q}{N_V} \epsilon \frac{\Delta L_t}{\Delta L_N} = \epsilon \frac{N_Q}{N_V} \tag{3}$$

since $\Delta L_t / \Delta L_N = 1$. The determination of the capacitance from a flux plot merely consists of counting squares in two directions, between conductors and around either conductor. From Fig. 6.3 we obtain

$$C = \epsilon_0 \frac{8 \times 3.25}{4} = 57.6 \quad \text{pF/m}$$

Ramo, Whinnery, and Van Duzer have an excellent discussion with examples of the construction of field maps by curvilinear squares. They offer the following suggestions:[1]

1. Plan on making a number of rough sketches, taking only a minute or so apiece, before starting any plot to be made with care. The use of transparent paper over the basic boundary will speed up this preliminary sketching.
2. Divide the known potential difference between electrodes into an equal number of divisions, say four or eight to begin with.

[1] By permission from S. Ramo, J. R. Whinnery, and T. Van Duzer, pp. 51–52. See Suggested References at the end of Chap. 5. Curvilinear maps are discussed on pp. 50–52.

3. Begin the sketch of equipotentials in the region where the field is known best, as for example in some region where it approaches a uniform field. Extend the equipotentials according to your best guess throughout the plot. Note that they should tend to hug acute angles of the conducting boundary and be spread out in the vicinity of obtuse angles of the boundary.
4. Draw in the orthogonal set of field lines. As these are started, they should form curvilinear squares, but, as they are extended, the condition of orthogonality should be kept paramount, even though this will result in some rectangles with ratios other than unity.
5. Look at the regions with poor side ratios and try to see what was wrong with the first guess of equipotentials. Correct them and repeat the procedure until reasonable curvilinear squares exist throughout the plot.
6. In regions of low field intensity, there will be large figures, often of five or six sides. To judge the correctness of the plot in this region, these large units should be subdivided. The subdivisions should be started back away from the region needing subdivision, and each time a flux tube is divided in half, the potential divisions in this region must be divided by the same factor.

✔ **D6.1.** Figure 6.4 shows the cross section of two circular cylinders at potentials of 0 and 60 V. The axes are parallel and the region between the cylinders is air-filled.

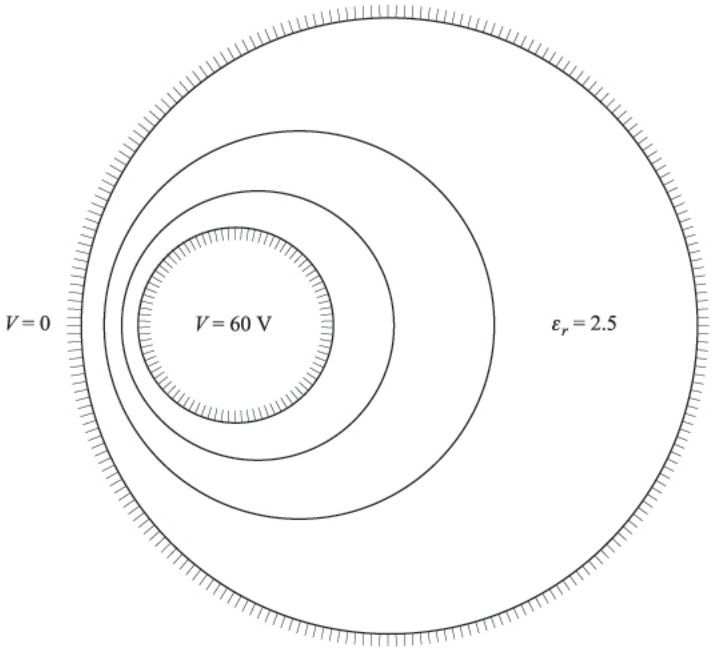


FIGURE 6.4
See Prob. D6.1.

Equipotentials at 20 V and 40 V are also shown. Prepare a curvilinear-square map on the figure and use it to establish suitable values for: (*a*) the capacitance per meter length; (*b*) E at the left side of the 60-V conductor if its true radius is 2 mm; (*c*) ρ_S at that point.

Ans. 69 pF/m; 60 kV/m; 550 nC/m^2

6.2 THE ITERATION METHOD

In potential problems where the potential is completely specified on the boundaries of a given region, particularly problems in which the potential does not vary in one direction (i.e., two-dimensional potential distributions) there exists a pencil-and-paper repetitive method which is capable of yielding any desired accuracy. Digital computers should be used when the value of the potential is required with high accuracy; otherwise, the time required is prohibitive except in the simplest problems. The iterative method, to be described below, is well suited for calculation by any digital computer.

Let us assume a two-dimensional problem in which the potential does not vary with the z coordinate and divide the interior of a cross section of the region where the potential is desired into squares of length h on a side. A portion of this region is shown in Fig. 6.5. The unknown values of the potential at five adjacent points are indicated as V_0, V_1, V_2, V_3, and V_4. If the region is charge-free and contains a homogeneous dielectric, then $\nabla \cdot \mathbf{D} = 0$ and $\nabla \cdot \mathbf{E} = 0$, from which we have, in two dimensions,

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0$$

But the gradient operation gives $E_x = -\partial V/\partial x$ and $E_y = -\partial V/\partial y$, from which we obtain[2]

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$$

Approximate values for these partial derivatives may be obtained in terms of the assumed potentials, or

$$\left.\frac{\partial V}{\partial x}\right|_a \doteq \frac{V_1 - V_0}{h}$$

and

$$\left.\frac{\partial V}{\partial x}\right|_c \doteq \frac{V_0 - V_3}{h}$$

[2] This is Laplace's equation in two dimensions. The three-dimensional form will be derived in the following chapter.

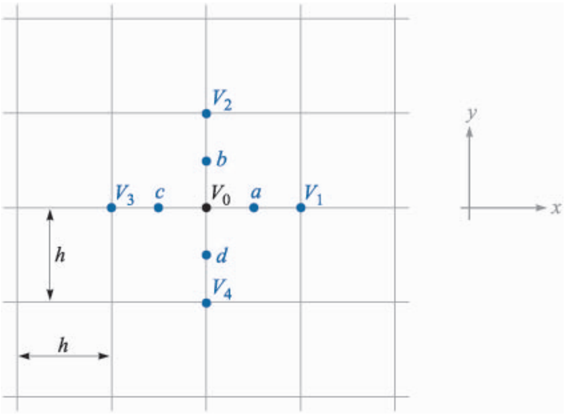


FIGURE 6.5
A portion of a region containing a two-dimensional potential field, divided into squares of side h. The potential V_0 is approximately equal to the average of the potentials at the four neighboring points.

from which

$$\left.\frac{\partial^2 V}{\partial x^2}\right|_0 \doteq \frac{\left.\frac{\partial V}{\partial x}\right|_a - \left.\frac{\partial V}{\partial x}\right|_c}{h} \doteq \frac{V_1 - V_0 - V_0 + V_3}{h^2}$$

and similarly,

$$\left.\frac{\partial^2 V}{\partial y^2}\right|_0 \doteq \frac{V_2 - V_0 - V_0 + V_4}{h^2}$$

Combining, we have

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} \doteq \frac{V_1 + V_2 + V_3 + V_4 - 4V_0}{h^2} = 0$$

or

$$\boxed{V_0 \doteq \tfrac{1}{4}(V_1 + V_2 + V_3 + V_4)} \tag{4}$$

The expression becomes exact as h approaches zero, and we shall write it without the approximation sign. It is intuitively correct, telling us that the potential is the average of the potential at the four neighboring points. The iterative method merely uses (4) to determine the potential at the corner of every square subdivision in turn, and then the process is repeated over the entire region as many times as is necessary until the values no longer change. The method is best shown in detail by an example.

For simplicity, consider a square region with conducting boundaries (Fig. 6.6). The potential of the top is 100 V and that of the sides and bottom is zero. The problem is two-dimensional, and the sketch is a cross section of the physical configuration. The region is divided first into 16 squares, and some estimate of

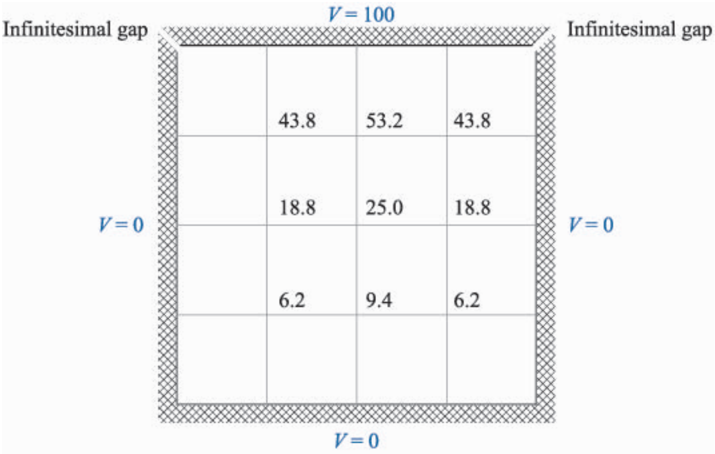


FIGURE 6.6
Cross section of a square trough with sides and bottom at zero potential and top at 100 V. The cross section has been divided into 16 squares, with the potential estimated at every corner. More accurate values may be determined by using the iteration method.

the potential must now be made at every corner before applying the iterative method. The better the estimate, the shorter the solution, although the final result is independent of these initial estimates. When the computer is used for iteration, the initial potentials are usually set equal to zero to simplify the program. Reasonably accurate values could be obtained from a rough curvilinear-square map, or we could apply (4) to the large squares. At the center of the figure the potential estimate is then $\frac{1}{4}(100 + 0 + 0 + 0) = 25.0$

The potential may now be estimated at the centers of the four double-sized squares by taking the average of the potentials at the four corners or applying (4) along a diagonal set of axes. Use of this "diagonal average" is made only in preparing initial estimates. For the two upper double squares, we select a potential of 50 V for the gap (the average of 0 and 100), and then $V = \frac{1}{4}(50 + 100 + 25 + 0) = 43.8$ (to the nearest tenth of a volt[3]), and for the lower ones,

$$V = \frac{1}{4}(0 + 25 + 0 + 0) = 6.2$$

The potential at the remaining four points may now be obtained by applying (4) directly. The complete set of estimated values is shown in Fig. 6.6.

The initial traverse is now made to obtain a corrected set of potentials, beginning in the upper left corner (with the 43.8 value, not with the boundary

[3] When rounding off a decimal ending exactly with a five, the preceding digit should be made *even* (e.g., 42.75 becomes 42.8 and 6.25 becomes 6.2). This generally ensures a random process leading to better accuracy than would be obtained by always increasing the previous digit by 1.

where the potentials are known and fixed), working across the row to the right, and then dropping down to the second row and proceeding from left to right again. Thus the 43.8 value changes to $\frac{1}{4}(100 + 53.2 + 18.8 + 0) = 43.0$. The best or newest potentials are always used when applying (4), so both points marked 43.8 are changed to 43.0, because of the evident symmetry, and the 53.2 value becomes $\frac{1}{4}(100 + 43.0 + 25.0 + 43.0) = 52.8$.

Because of the symmetry, little would be gained by continuing across the top line. Each point of this line has now been improved once. Dropping down to the next line, the 18.8 value becomes

$$V = \frac{1}{4}(43.0 + 25.0 + 6.2 + 0) = 18.6$$

and the traverse continues in this manner. The values at the end of this traverse are shown as the top numbers in each column of Fig. 6.7. Additional traverses must now be made until the value at each corner shows no change. The values for the successive traverses are usually entered below each other in column form, as shown in Fig. 6.7, and the final value is shown at the bottom of each column. Only four traverses are required in this example.

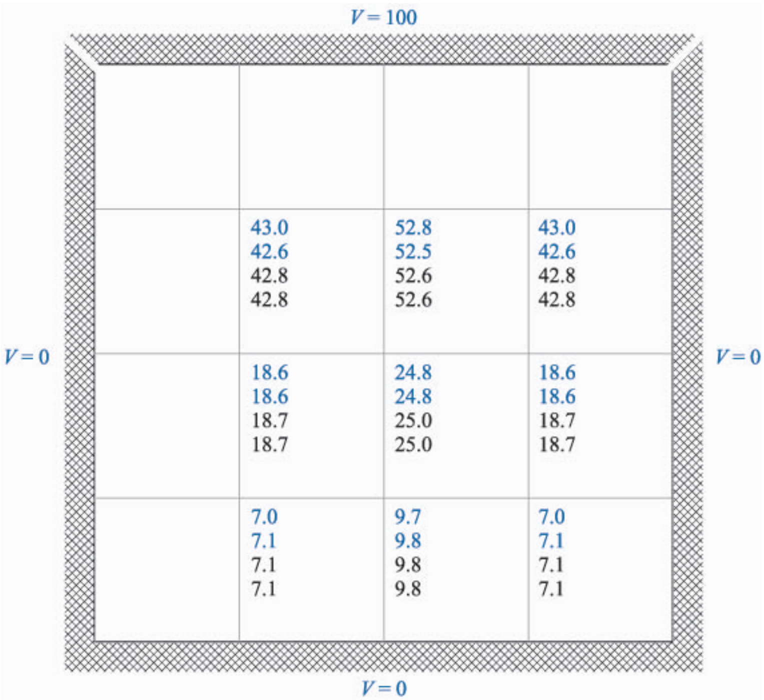


FIGURE 6.7
The results of each of the four necessary traverses of the problem of Fig. 6.5 are shown in order in the columns. The final values, unchanged in the last traverse, are at the bottom of each column.

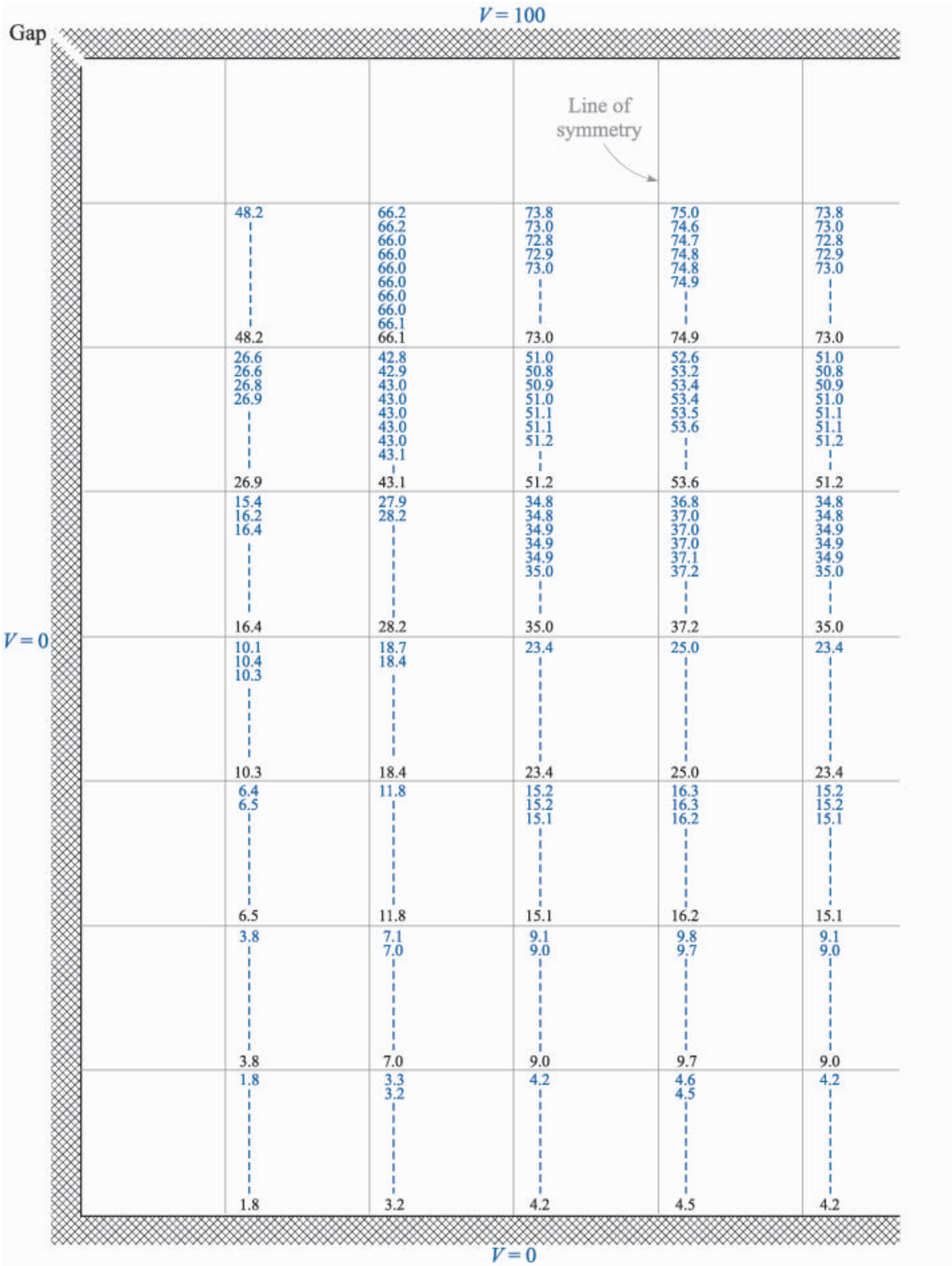


FIGURE 6.8
The problem of Figs. 6.6 and 6.7 is divided into smaller squares. Values obtained on the nine successive traverses are listed in order in the columns.

If each of the nine initial values were set equal to zero, it is interesting to note that ten traverses would be required. The cost of having a computer do these additional traverses is probably much less than the cost of the programming necessary to make decent initial estimates.

Since there is a large difference in potential from square to square, we should not expect our answers to be accurate to the tenth of a volt shown (and perhaps not to the nearest volt). Increased accuracy comes from dividing each square into four smaller squares, not from finding the potential to a larger number of significant figures at each corner.

In Fig. 6.8, which shows only one of the symmetrical halves plus an additional column, this subdivision is accomplished, and the potential at the newly created corners is estimated by applying (4) directly where possible and diagonally when necessary. The set of estimated values appears at the top of each column, and the values produced by the successive traverses appear in order below. Here nine sets of values are required, and it might be noted that no values change on the last traverse (a necessary condition for the *last* traverse), and only one value changes on each of the preceding three traverses. No value in the bottom four rows changes after the second traverse; this results in a great saving in time, for if none of the four potentials in (4) changes, the answer is of course unchanged.

For this problem, it is possible to compare our final values with the exact potentials, obtained by evaluating some infinite series, as discussed at the end of the following chapter. At the point for which the original estimate was 53.2, the final value for the coarse grid was 52.6, the final value for the finer grid was 53.6, and the final value for a 16×16 grid is 53.93 V to two decimals, using data obtained with the following Fortran program:

```
 1  DIMENSION A (17,17),B(17,17)
 2  DO 6 I=2,17
 3  DO 5 J=1,17
 4  A(I,J)=0.
 5  CONTINUE
 6  CONTINUE
 7  DO 9 J=2,16
 8  A(I,J)=100.
 9  CONTINUE
10  A(1,1)=50.
11  A(1,17)=50.
12  DO 16 I=2,16
13  DO 15 J=2,16
14  A(I,J)=(A(I,J-1)+A(I-1,J)+A(I,J+1)+A(I+1,J))/4.
15  CONTINUE
16  CONTINUE
17  DO 23 I=2,16
18  DO 22 J=2,16
```

```
19   C=(A(I,J-1)+A(I-1,J)+A(I,J+1)+A(I+1,J))/4.
20   B(I,J)=A(I,J)-C
21   IF((ABS(B(I,J))-.00001).GT.0.) GO TO 12
22   CONTINUE
23   CONTINUE
24   WRITE(6,25)((A(I,J),J=1,17),I=1,17)
25   FORMAT (1H0,17F7.2)
26   STOP
27   END
```

Line 21 shows that the iteration is continued until the difference between successive traverses is less than 10^{-5}.

The exact potential obtained by a Fourier expansion is 54.05 V to two decimals. Two other points are also compared in tabular form, as shown in Table 6.1.

Computer flowcharts and programs for iteration solutions are given in Chap. 24 of Boast[4] and in Chap. 2 and the appendix of Silvester.[5]

Very few electrode configurations have a square or rectangular cross section that can be neatly subdivided into a square grid. Curved boundaries, acute- or obtuse-angled corners, reentrant shapes, and other irregularities require slight modifications of the basic method. An important one of these is described in Prob. 10 at the end of this chapter, and other irregular examples appear as Probs. 7, 9, and 11.

A refinement of the iteration method is known as the *relaxation method*. In general it requires less work but more care in carrying out the arithmetical steps.[6]

✔ **D6.2.** In Fig. 6.9, a square grid is shown within an irregular potential trough. Using the iteration method to find the potential to the nearest volt, determine the final value at: (*a*) point *a*; (*b*) point *b*; (*c*) point *c*.

Ans. 18 V; 46 V; 91 V

TABLE 6.1

Original estimate	53.2	25.0	9.4
4 × 4 grid	52.6	25.0	9.8
8 × 8 grid	53.6	25.0	9.7
16 × 16 grid	53.93	25.00	9.56
Exact	54.05	25.00	9.54

[4]See Suggested References at the end of Chap. 2.

[5] See Suggested References at the end of this chapter.

[6] A detailed description appears in Scarborough, and the basic procedure and one example are in an earlier edition of Hayt. See Suggested References at the end of the chapter.

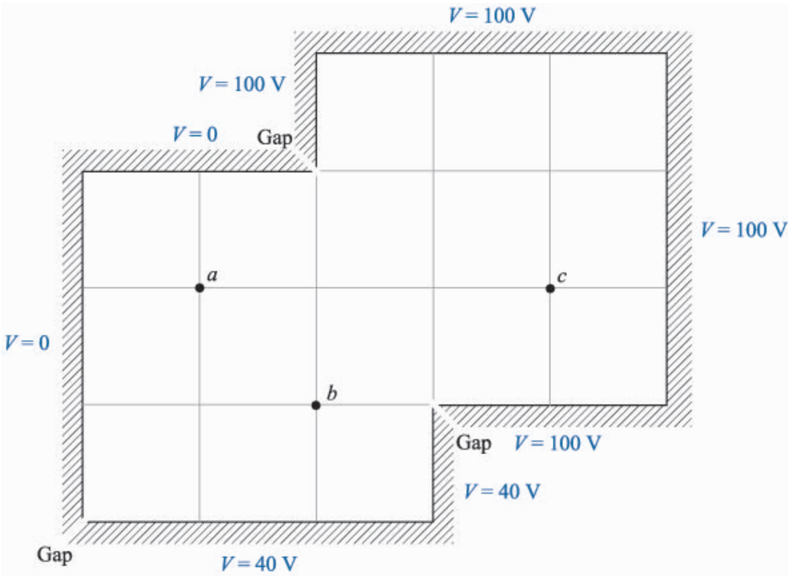


FIGURE 6.9
See Prob. D6.2.

6.3 CURRENT ANALOGIES

Several experimental methods depend upon an analogy between current density in conducting media and electric flux density in dielectric media. The analogy is easily demonstrated, for in a conducting medium Ohm's law and the gradient relationship are, for direct currents only,

$$\mathbf{J} = \sigma \mathbf{E}_\sigma$$

$$\mathbf{E}_\sigma = -\nabla V_\sigma$$

whereas in a homogeneous dielectric

$$\mathbf{D} = \epsilon \mathbf{E}_\epsilon$$

$$\mathbf{E}_\epsilon = -\nabla V_\epsilon$$

The subscripts serve to identify the analogous problems. It is evident that the potentials V_σ and V_ϵ, the electric field intensities $\mathbf{E}_\sigma$ and $\mathbf{E}_\epsilon$, the conductivity and permittivity σ and ϵ, and the current density and electric flux density $\mathbf{J}$ and $\mathbf{D}$ are analogous in pairs. Referring to a curvilinear-square map, we should interpret flux tubes as current tubes, each tube now carrying an incremental current which cannot leave the tube.

Finally, we must look at the boundaries. What is analogous to a conducting boundary which terminates electric flux normally and is an equipotential surface? The analogy furnishes the answer, and we see that the surface must termi-

nate current density normally and again be an equipotential surface. This is the surface of a *perfect* conductor, although in practice it is necessary only to use one whose conductivity is many times that of the conducting medium.

Therefore, if we wished to find the field within a coaxial capacitor, which, as we have seen several times before, is a portion of the field of an infinite line charge, we might take two copper cylinders and fill the region between them with, for convenience, an electrolytic solution. Applying a potential difference betwen the cylinders, we may use a probe to establish the potential at any intermediate point, or to find all those points having the same potential. This is the essence of the electrolytic trough or tank. The greatest advantage of this method lies in the fact that it is not limited to two-dimensional problems. Practical suggestions for the construction and use of the tank are given in many places.[7]

The determination of capacitance from electrolytic-trough measurements is particularly easy. The total current leaving the more positive conductor is

$$I = \oint_S \mathbf{J} \cdot d\mathbf{S} = \sigma \oint_S \mathbf{E}_\sigma \cdot d\mathbf{S}$$

where the closed surface integral is taken over the entire conductor surface. The potential difference is given by the negative line integral from the less to the more positive plate,

$$V_{\sigma 0} = -\int \mathbf{E}_\sigma \cdot d\mathbf{L}$$

and the total resistance is therefore

$$R = \frac{V_{\sigma 0}}{I} = \frac{-\int \mathbf{E}_\sigma \cdot d\mathbf{L}}{\sigma \oint_S \mathbf{E}_\sigma \cdot d\mathbf{S}}$$

The capacitance is given by the ratio of the total charge to the potential difference,

$$C = \frac{Q}{V_{\epsilon 0}} = \frac{\epsilon \oint_S \mathbf{E}_\epsilon \cdot d\mathbf{S}}{-\int \mathbf{E}_\epsilon \cdot d\mathbf{L}}$$

We now invoke the analogy by letting $V_{\epsilon 0} = V_{\sigma 0}$ and $\mathbf{E}_\epsilon = \mathbf{E}_\sigma$. The result is

$$\boxed{RC = \frac{\epsilon}{\sigma}} \tag{5}$$

Knowing the conductivity of the electrolyte and the permittivity of the dielectric, we may determine the capacitance by a simple resistance measurement.

A simpler technique is available for two-dimensional problems. Conducting paper is used as the base on which the conducting boundaries are drawn with

[7] Weber is good. See Suggested References at the end of the chapter.

silver paint. In the case of the coaxial capacitor, we should draw two circles of radii ρ_A and ρ_B, $\rho_B > \rho_A$, extending the paint a small distance outward from ρ_B and inward from ρ_A to provide sufficient area to make a good contact with wires to an external potential source. A probe is again used to establish potential values between the circles.

Conducting paper is described in terms of its *sheet resistance* R_S. The sheet resistance is the resistance between opposite edges of a square. Since the fields are uniform in such a square, we may apply Eq. (13) from Chap. 5,

$$R = \frac{L}{\sigma S}$$

to the case of a square of conducting paper having a width w and a thickness t, where $w = L$,

$$R_S = \frac{L}{\sigma t L} = \frac{l}{\sigma t} \quad \Omega \tag{6}$$

Thus, if the conductive coating has a thickness of 0.2 mm and a conductivity of 2 S/m, its surface resistance is $1/(2 \times 0.2 \times 10^{-3}) = 2500\,\Omega$. The units of R_S are often given as "ohms per square" (but never as ohms per square meter).

Fig. 6.10 shows the silver-paint boundaries that would be drawn on the conducting paper to determine the capacitance of a square-in-a-circle transmission line like that shown in Fig. 6.3. The generator and detector often operate at 1 kHz to permit use of a more sensitive tuned detector or bridge.

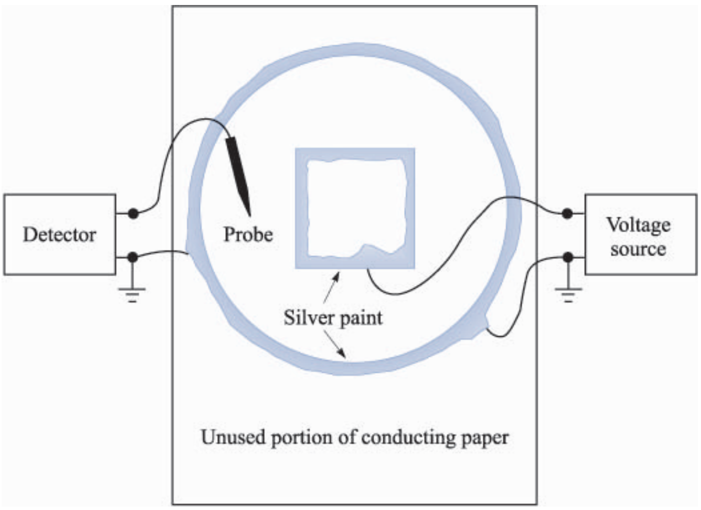


FIGURE 6.10
A two-dimensional, two-conductor problem similar to that of Fig. 6.3 is drawn on conducting paper. The probe may be used to trace out an equipotential surface.

✔ **D6.3.** If the conducting paper shown in Fig. 6.10 has a sheet resistance of 1800 Ω per square, find the resistance that would be measured between the opposite edges of: (*a*) a square 9 cm on a side; (*b*) a square 4.5 cm on a side; (*c*) a rectangle 3 cm by 9 cm, across the longer dimension; (*d*) a rectangle 3 cm by 9 cm, across the shorter dimension. (*e*) What resistance would be measured between an inner circle of 0.8-cm radius and an outer circle of 2-cm radius?

Ans**.** 1800 Ω; 1800 Ω; 5400 Ω; 600 Ω; 397 Ω

6.4 PHYSICAL MODELS

The analogy between the electric field and the gravitational field was mentioned several times previously and may be used to construct physical models which are capable of yielding solutions to electrostatic problems of complicated geometry. The basis of the analogy is simply this: in the electrostatic field the potential difference between two points is the difference in the potential energy of unit positive charges at these points, and in a uniform gravitational field the difference in the potential energy of point masses at two points is proportional to their difference in height. In other words,

$$\Delta W_E = Q\Delta V \quad \text{(electrostatic)}$$
$$\Delta W_G = Mg\Delta h \quad \text{(gravitational)}$$

where M is the point mass and g is the acceleration due to gravity, essentially constant at the surface of the earth. For the same energy difference, then,

$$\Delta V = \frac{Mg}{Q}\Delta h = k\Delta h$$

where k is the constant of proportionality. This shows the direct analogy between difference in potential and difference in elevation.

This analogy allows us to construct a physical model of a known two-dimensional potential field by fabricating a surface, perhaps from wood, whose elevation h above any point (x, y) located in the zero-elevation zero-potential plane is proportional to the potential at that point. Note that three-dimensional fields cannot be handled.

The field of an infinite line charge,

$$V = \frac{\rho_L}{2\pi\epsilon}\ln\frac{\rho_B}{\rho}$$

is shown on such a model in Fig. 6.11, which provides an accurate picture of the variation of potential with radius between ρ_A and ρ_B. The potential and elevation at ρ_B are conveniently set equal to zero.

Such a model may be constructed for any two-dimensional potential field and enables us to visualize the field a little better. The construction of the models themselves is enormously simplified, both physically and theoretically, by the use of rubber sheets. The sheet is placed under moderate tension and approximates

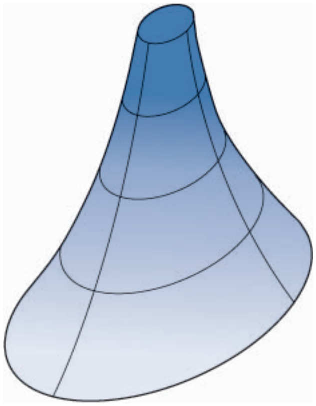


FIGURE 6.11
A model of the potential field of an infinite line charge. The difference in potential is proportional to the difference in elevation. Contour lines indicate equal potential increments.

closely the *elastic membrane* of applied mechanics. It can be shown[8] that the vertical displacement h of the membrane satisfies the second-order partial differential equation

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0$$

if the surface slope is small. We shall see in the next chapter that every potential field in a charge-free region also satisfies this equation, Laplace's equation in two dimensions,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$$

We shall also prove a uniqueness theorem which assures us that if a potential solution in some specified region satisfies the above equation and also gives the correct potential on the boundaries of this region, then this solution is the only solution. Hence we need only force the elevation of the sheet to corresponding prescribed potential values on the boundaries, and the elevation at all other points is proportional to the potential.

For instance, the infinite-line-charge field may be displayed by recognizing the circular symmetry and fastening the rubber sheet at zero elevation around a circle by the use of a large clamping ring of radius ρ_B. Since the potential is constant at ρ_A, we raise that portion of the sheet to a greater elevation by pushing a cylinder of radius ρ_A up against the rubber sheet. The analogy breaks down for large surface slopes, and only a slight displacement at ρ_A is possible. The surface then represents the potential field, and marbles may be used to determine particle trajectories, in this case obviously radial lines as viewed from above.

[8] See, for instance, Spangenberg, pp. 75–76, in Suggested References at the end of the chapter.

There is also an analogy between electrostatics and hydraulics that is particularly useful in obtaining photographs of the streamlines or flux lines. This process is described completely by Moore in a number of publications[9] which include many excellent photographs.

✔ **D6.4.** A potential field, $V = 200(x^2 - 4y + 2)$ V, is illustrated by a plaster model with a scale of 1 vertical inch = 400 V; the horizontal dimensions are true. The region shown is $3 \le x \le 4$ m, $0 \le y \le 1$ m. In this region: (*a*) What is the maximum height of the model? (*b*) What is its minimum height? (*c*) What is the difference in height between points $A(x = 3.2, y = 0.5)$ and $B(3.8, 1)$? (*d*) What angle does the line connecting these two points make with the horizontal?

SUGGESTED REFERENCES

1. Hayt, W. H., Jr.: "Engineering Electromagnetics," 1st ed., McGraw-Hill Book Company, New York, 1958, pp. 150–152.
2. Moore, A. D.: Fields from Fluid Flow Mappers, *J. Appl. Phys.*, vol. 20, pp. 790-804, August 1949; Soap Film and Sandbed Mapper Techniques, *J. Appl. Mech.* (bound with *Trans. ASME*), vol. 17, pp. 291–298, September 1950; Four Electromagnetic Propositions, with Fluid Mapper Verifications, *Elec. Eng.*, vol. 69, pp. 607–610, July 1950; The Further Development of Fluid Mappers, *Trans. AIEE*, vol. 69, part II, pp. 1615–1624, 1950; Mapping Techniques Applied to Fluid Mapper Patterns, *Trans. AIEE*, vol. 71, part I, pp. 1–5, 1952.
3. Salvadori, M. G., and M. L. Baron: "Numerical Methods in Engineering," 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961. Iteration and relaxation methods are discussed in chap. 1.
4. Scarborough, J. B.: "Numerical Mathematical Analysis," 6th ed., The Johns Hopkins Press, Baltimore, 1966. Describes iteration relaxation methods and gives several complete examples. Inherent errors are discussed.
5. Silvester, P.: "Modern Electromagnetic Fields," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968.
6. Soroka, W. W.: "Analog Methods in Computation and Simulation," McGraw-Hill Book Company, New York, 1954.
7. Spangenberg, K. R.: "Vacuum Tubes," McGraw-Hill Book Company, New York, 1948. Experimental mapping methods are discussed on pp. 75–82.
8. Weber, E.: "Electromagnetic Fields," vol. I, John Wiley & Sons, Inc., New York, 1950. Experimental mapping methods are discussed in chap. 5.

[9] See Suggested References at the end of this chapter.

PROBLEMS

6.1 Construct a curvilinear-square map for a coaxial capacitor of 3-cm inner radius and 8-cm outer radius. These dimensions are suitable for the drawing. (*a*) Use your sketch to calculate the capacitance per meter length, assuming $\epsilon_R = 1$. (*b*) Calculate an exact value for the capacitance per unit length.

6.2 Construct a curvilinear-square map of the potential field about two parallel circular cylinders, each of 2.5-cm radius, separated a center-to-center distance of 13 cm. These dimensions are suitable for the actual sketch if symmetry is considered. As a check, compute the capacitance per meter both from your sketch and from the exact formula. Assume $\epsilon_R = 1$.

6.3 Construct a curvilinear-square map of the potential field between two parallel circular cylinders, one of 4-cm radius inside one of 8-cm radius. The two axes are displaced by 2.5 cm. These dimensions are suitable for the drawing. As a check on the accuracy, compute the capacitance per meter from the sketch and from the exact expression:

$$\frac{2\pi\epsilon}{\cosh^{-1}\dfrac{a^2 + b^2 - D^2}{2ab}}$$

where a and b are the conductor radii and D is the axis separation.

6.4 A solid conducting cylinder of 4-cm radius is centered within a rectangular conducting cylinder with a 12-cm by 20-cm cross section. (*a*) Make a full-size sketch of one quadrant of this configuration and construct a curvilinear-square map for its interior. (*b*) Assume $\epsilon = \epsilon_0$ and estimate C per meter length.

6.5 The inner conductor of the transmission line shown in Fig. 6.12 has a square cross section $2a \times 2a$, while the outer square is $5a \times 5a$. The axes are displaced as shown. (*a*) Construct a good-sized drawing of this transmission line, say with $a = 2.5$ cm, and then prepare a curvilinear-square plot of the electrostatic field between the conductors. (*b*) Use your map to calculate the capacitance per meter length if $\epsilon = 1.6\epsilon_0$. (*c*) How would the answer to part *b* change if $a = 0.6$ cm?

6.6 Let the inner conductor of the transmission line shown in Fig. 6.12 be at a potential of 100 V, while the outer is at zero potential. Construct a grid, $0.5a$ on a side, and use iteration to find V at a point that is a units above the upper right corner of the inner conductor. Work to the nearest volt.

6.7 Use the iteration method to estimate the potential at points x and y in the triangular trough of Fig. 6.13. Work only to the nearest volt.

6.8 Use iteration methods to estimate the potential at point x in the trough shown in Fig. 6.14. Working to the nearest volt is sufficient.

6.9 Using the grid indicated in Fig. 6.15, work to the nearest volt to estimate the potential at point A.

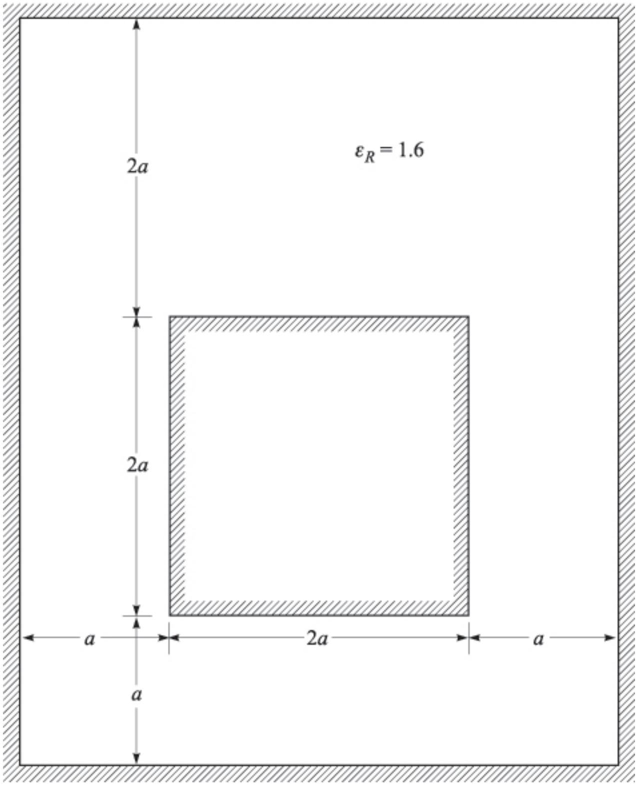


FIGURE 6.12
See Probs. 5, 6, and 14.

6.10 Conductors having boundaries that are curved or skewed usually do not permit every grid point to coincide with the actual boundary. Figure 6.16*a* illustrates the situation where the potential at V_0 is to be estimated in terms of V_1, V_2, V_3, V_4, and the unequal distances h_1, h_2, h_3, and h_4. (*a*) Show that

$$V_0 = \frac{V_1}{\left(1+\frac{h_1}{h_3}\right)\left(1+\frac{h_1 h_3}{h_4 h_2}\right)} + \frac{V_2}{\left(1+\frac{h_2}{h_4}\right)\left(1+\frac{h_2 h_4}{h_4 h_3}\right)} + \frac{V_3}{\left(1+\frac{h_3}{h_1}\right)\left(1+\frac{h_3 h_1}{h_2 h_4}\right)}$$

$$+ \frac{V_4}{\left(1+\frac{h_4}{h_2}\right)\left(1+\frac{h_4 h_2}{h_3 h_1}\right)};$$ (*b*) determine V_0 in Fig. 6.16*b*.

6.11 Consider the configuration of conductors and potentials shown in Fig. 6.17. Using the method described in Prob. 10, write an expression for V_0

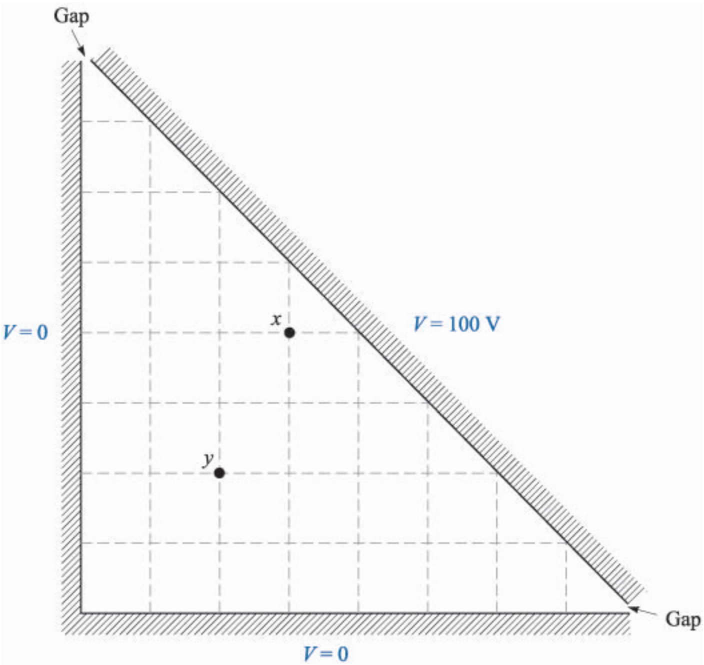


FIGURE 6.13
See Prob. 7.

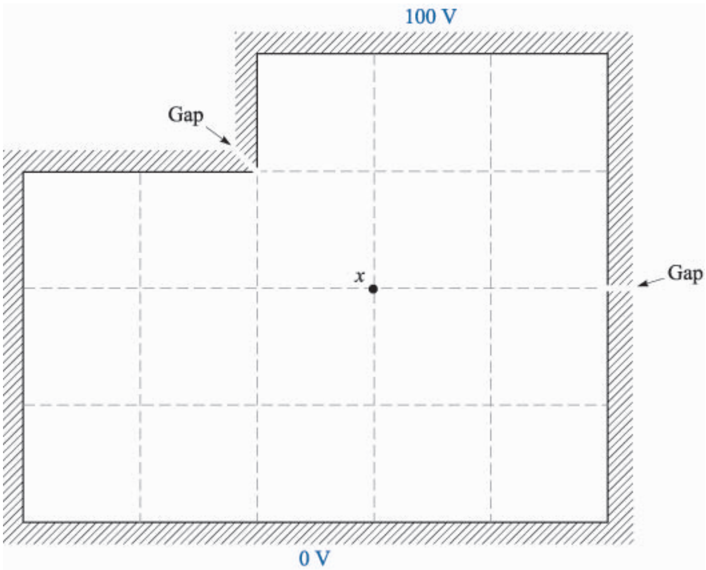


FIGURE 6.14
See Prob. 8.

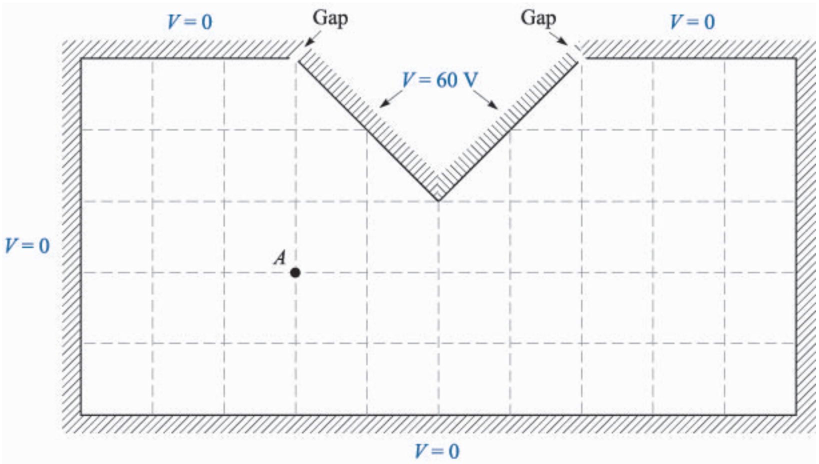


FIGURE 6.15
See Prob. 9.

in terms of V_1, V_2, V_3, and V_4 at point: (*a*) x; (*b*) y; (*c*) z. (*d*) Use the iteration method to estimate the potential at point x.

6.12 (*a*) After estimating potentials for the configuration of Fig. 6.18, use the iteration method with a square grid 1 cm on a side to find better estimates at the seven grid points. Work to the nearest volt. (*b*) Construct 0.5-cm grid, establish new rough estimates, and then use the iteration method on the 0.5-cm grid. Again work to the nearest volt. (*c*) Use the computer to obtain values for a 0.25-cm grid. Work to the nearest 0.1 V.

6.13 Perfectly conducting concentric spheres have radii of 2 and 6 cm. The region $2 < r < 3$ cm is filled with a solid conducting material for which

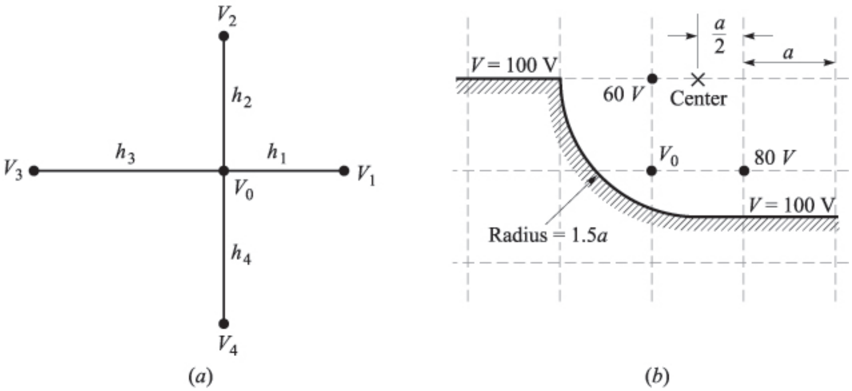


FIGURE 6.16
See Prob. 10.

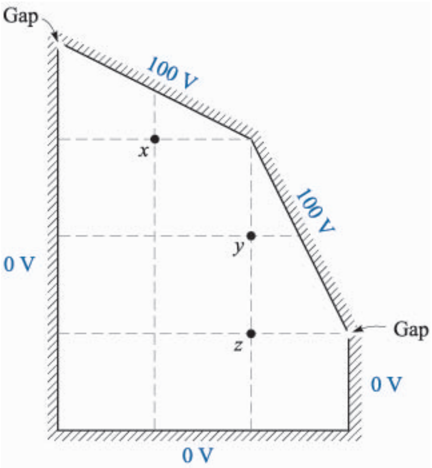


FIGURE 6.17
See Prob. 11.

$\sigma = 100\,\text{S/m}$, while the portion for which $3 < r < 6\,\text{cm}$ has $\sigma = 25\,\text{S/m}$. The inner sphere is held at 1 V while the outer is at $V = 0$. (*a*) Find E and J everywhere. (*b*) What resistance would be measured between the two spheres? (*c*) What is V at $r = 3\,\text{cm}$?

6.14 The cross section of the transmission line shown in Fig. 6.12 is drawn on a sheet of conducting paper with metallic paint. The sheet resistance is $2000\,\Omega$ per square and the dimension a is 2 cm. (*a*) Assuming a result for Prob. 6*b* of 110 pF/m, what total resistance would be measured between the metallic conductors drawn on the conducting paper? (*b*) What would the total resistance be if $a = 2\,\text{cm}$?

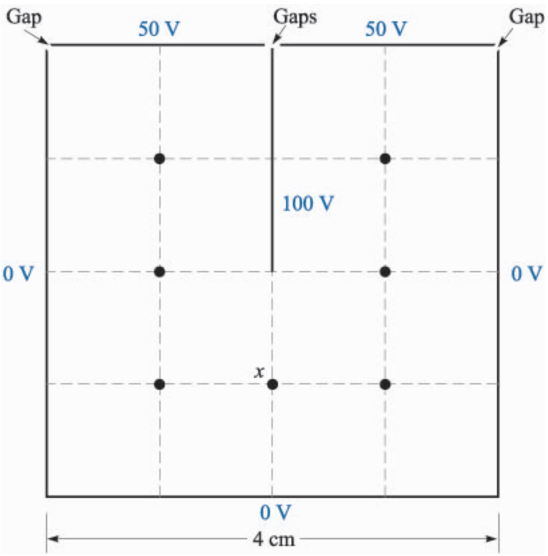


FIGURE 6.18
See Prob. 12.

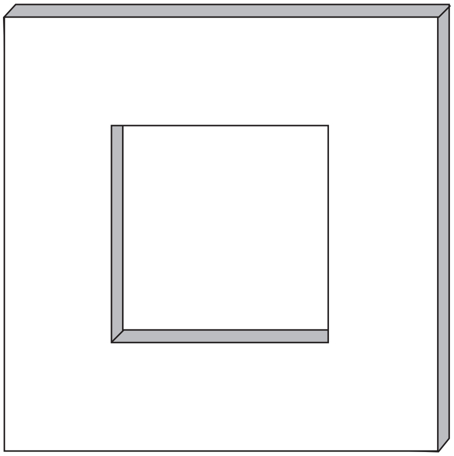


FIGURE 6.19
See Prob. 16.

6.15 Two concentric annular rings are painted on a sheet of conducting paper with a highly conducting metallic paint. The four radii are 1, 1.2, 3.5, and 3.7 cm. Connections made to the two rings show a resistance of 215 Ω between them. (*a*) What is R_S for the conducting paper? (*b*) If the conductivity of the material used as the surface of the paper is 2 S/m, what is the thickness of the coating?

6.16 The square washer shown in Fig. 6.19 is 2.4 mm thick and has outer dimensions of 2.5×2.5 cm and inner dimensions of 1.25×1.25 cm. The inside and outside surfaces are perfectly conducting. If the material has a conductivity of 6 S/m, estimate the resistance offered between the inner and outer surfaces (shown shaded in Fig. 6.19). A few curvilinear squares are suggested.

6.17 A two-wire transmission line consists of two parallel perfectly conducting cylinders, each having a radius of 0.2 mm, separated a center-to-center distance of 2 mm. The medium surrounding the wires has $\epsilon_R = 3$ and $\sigma = 1.5$ mS/m. A 100-V battery is connected between the wires. Calculate: (*a*) the magnitude of the charge per meter length on each wire; (*b*) the battery current.

6.18 A coaxial transmission line is modelled by the use of a rubber sheet having horizontal dimensions that are 100 times those of the actual line. Let the radial coordinate of the model be ρ_m. For the line itself, let the radial dimension be designated by ρ as usual; also, let $a = 0.6$ mm and $b = 4.8$ mm. The model is 8 cm in height at the inner conductor and zero at the outer. If the potential of the inner conductor is 100 V: (*a*) find the expression for $V(\rho)$. (*b*) Write the model height as a function of ρ.

CHAPTER
7

POISSON'S AND LAPLACE'S EQUATIONS

A study of the previous chapter shows that several of the analogies used to obtain experimental field maps involved demonstrating that the analogous quantity satisfies Laplace's equation. This is true for small deflections of an elastic membrane, and we might have proved the current analogy by showing that the direct-current density in a conducting medium also satisfies Laplace's equation. It appears that this is a fundamental equation in more than one field of science, and, perhaps without knowing it, we have spent the last chapter obtaining solutions for Laplace's equation by experimental, graphical, and numerical methods. Now we are ready to obtain this equation formally and discuss several methods by which it may be solved analytically.

It may seem that this material properly belongs before that of the previous chapter; as long as we are solving one equation by so many methods, would it not be fitting to see the equation first? The disadvantage of this more logical order lies in the fact that solving Laplace's equation is an exercise in mathematics, and unless we have the physical problem well in mind, we may easily miss the physical significance of what we are doing. A rough curvilinear map can tell us much about a field and then may be used later to check our mathematical solutions for gross errors or to indicate certain peculiar regions in the field which require special treatment.

With this explanation let us finally obtain the equations of Laplace and Poisson.

7.1 POISSON'S AND LAPLACE'S EQUATIONS

Obtaining Poisson's equation is exceedingly simple, for from the point form of Gauss's law,

$$\nabla \cdot \mathbf{D} = \rho_v \tag{1}$$

the definition of **D**,

$$\mathbf{D} = \epsilon \mathbf{E} \tag{2}$$

and the gradient relationship,

$$\mathbf{E} = -\nabla V \tag{3}$$

by substitution we have

$$\nabla \cdot \mathbf{D} = \nabla \cdot (\epsilon \mathbf{E}) = -\nabla \cdot (\epsilon \nabla V) = \rho_v$$

or

$$\nabla \cdot \nabla V = -\frac{\rho_v}{\epsilon} \tag{4}$$

for a homogeneous region in which ϵ is constant.

Equation (4) is *Poisson's equation*, but the "double ∇" operation must be interpreted and expanded, at least in cartesian coordinates, before the equation can be useful. In cartesian coordinates,

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

$$\nabla V = \frac{\partial V}{\partial x}\mathbf{a}_x + \frac{\partial V}{\partial y}\mathbf{a}_y + \frac{\partial V}{\partial z}\mathbf{a}_z$$

and therefore

$$\nabla \cdot \nabla V = \frac{\partial}{\partial x}\left(\frac{\partial V}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial V}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial V}{\partial z}\right)$$

$$= \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \tag{5}$$

Usually the operation $\nabla \cdot \nabla$ is abbreviated ∇^2 (and pronounced "del squared"), a good reminder of the second-order partial derivatives appearing in (5), and we have

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\rho_v}{\epsilon} \tag{6}$$

in cartesian coordinates.

If $\rho_v = 0$, indicating zero *volume* charge density, but allowing point charges, line charge, and surface charge density to exist at singular locations as sources of the field, then

$$\nabla^2 V = 0 \tag{7}$$

which is *Laplace's equation*. The ∇^2 operation is called the *Laplacian of V*.

In cartesian coordinates Laplace's equation is

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \qquad \text{(cartesian)} \tag{8}$$

and the form of $\nabla^2 V$ in cylindrical and spherical coordinates may be obtained by using the expressions for the divergence and gradient already obtained in those coordinate systems. For reference, the Laplacian in cylindrical coordinates is

$$\nabla^2 V = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial V}{\partial \rho}\right) + \frac{1}{\rho^2}\left(\frac{\partial^2 V}{\partial \phi^2}\right) + \frac{\partial^2 V}{\partial z^2} \qquad \text{(cylindrical)} \tag{9}$$

and in spherical coordinates is

$$\nabla^2 V = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial V}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 V}{\partial \phi^2} \qquad \text{(spherical)} \tag{10}$$

These equations may be expanded by taking the indicated partial derivatives, but it is usually more helpful to have them in the forms given above; furthermore, it is much easier to expand them later if necessary than it is to put the broken pieces back together again.

Laplace's equation is all-embracing, for, applying as it does wherever volume charge density is zero, it states that every conceivable configuration of electrodes or conductors produces a field for which $\nabla^2 V = 0$. All these fields are different, with different potential values and different spatial rates of change, yet for each of them $\nabla^2 V = 0$. Since *every* field (if $\rho_v = 0$) satisfies Laplace's equation, how can we expect to reverse the procedure and use Laplace's equation to find one specific field in which we happen to have an interest? Obviously, more

information is required, and we shall find that we must solve Laplace's equation subject to certain *boundary conditions*.

Every physical problem must contain at least one conducting boundary and usually contains two or more. The potentials on these boundaries are assigned values, perhaps V_0, $V_1, \ldots$, or perhaps numerical values. These definite equipotential surfaces will provide the boundary conditions for the type of problem to be solved in this chapter. In other types of problems, the boundary conditions take the form of specified values of E on an enclosing surface, or a mixture of known values of V and E.

Before using Laplace's equation or Poisson's equation in several examples, we must pause to show that if our answer satisfies Laplace's equation and also satisfies the boundary conditions, then it is the only possible answer. It would be very distressing to work a problem by solving Laplace's equation with two different approved methods and then to obtain two different answers. We shall show that the two answers must be identical.

✔ **D7.1.** Calculate numerical values for V and ρ_v at point P in free space if: (a) $V = \dfrac{4yz}{x^2+1}$, at $P(1,2,3)$; (b) $V = 5\rho^2 \cos 2\phi$, at $P(\rho = 3, \phi = \dfrac{\pi}{3}, z = 2)$; ($c$) $V = \dfrac{2\cos\phi}{r^2}$, at $P(r = 0.5, \theta = 45^\circ, \phi = 60^\circ)$.

Ans. 12 V, −106.2 pC/m^3; 22.5 V, 0; 4 V, −141.7 pC/m^3

7.2 UNIQUENESS THEOREM

Let us assume that we have two solutions of Laplace's equation, V_1 and V_2, both general functions of the coordinates used. Therefore

$$\nabla^2 V_1 = 0$$

and

$$\nabla^2 V_2 = 0$$

from which

$$\nabla^2 (V_1 - V_2) = 0$$

Each solution must also satisfy the boundary conditions, and if we represent the given potential values on the boundaries by V_b, then the value of V_1 on the boundary V_{1b} and the value of V_2 on the boundary V_{2b} must both be identical to V_b,

$$V_{1b} = V_{2b} = V_b$$

or

$$V_{1b} - V_{2b} = 0$$

In Sec. 4.8, Eq. (44), we made use of a vector identity,

$$\nabla \cdot (V\mathbf{D}) \equiv V(\nabla \cdot \mathbf{D}) + \mathbf{D} \cdot (\nabla V)$$

which holds for any scalar V and any vector $\mathbf{D}$. For the present application we shall select $V_1 - V_2$ as the scalar and $\nabla(V_1 - V_2)$ as the vector, giving

$$\nabla \cdot [(V_1 - V_2)\nabla(V_1 - V_2)] \equiv (V_1 - V_2)[\nabla \cdot \nabla(V_1 - V_2)] + \nabla(V_1 - V_2) \cdot \nabla(V_1 - V_2)$$

which we shall integrate throughout the volume *enclosed* by the boundary surfaces specified:

$$\int_{\text{vol}} \nabla \cdot [(V_1 - V_2)\nabla(V_1 - V_2)]\, dv \equiv \int_{\text{vol}} (V_1 - V_2)[\nabla \cdot \nabla(V_1 - V_2)] dv + \int_{\text{vol}} [\nabla(V_1 - V_2)]^2\, dv \qquad (11)$$

The divergence theorem allows us to replace the volume integral on the left side of the equation by the closed surface integral over the surface surrounding the volume. This surface consists of the boundaries already specified on which $V_{1b} = V_{2b}$, and therefore

$$\int_{\text{vol}} \nabla \cdot [(V_1 - V_2)\nabla(V_1 - V_2)]\, dv = \oint_S [(V_{1b} - V_{2b})\nabla(V_{1b} - V_{2b})] \cdot d\mathbf{S} = 0$$

One of the factors of the first integral on the right side of (11) is $\nabla \cdot \nabla(V_1 - V_2)$, or $\nabla^2(V_1 - V_2)$, which is zero by hypothesis, and therefore that integral is zero. Hence the remaining volume integral must be zero:

$$\int_{\text{vol}} [\nabla(V_1 - V_2)]^2\, dv = 0$$

There are two reasons why an integral may be zero: either the integrand (the quantity under the integral sign) is everywhere zero, or the integrand is positive in some regions and negative in others, and the contributions cancel algebraically. In this case the first reason must hold because $[\nabla(V_1 - V_2)]^2$ cannot be negative. Therefore

$$[\nabla(V_1 - V_2)]^2 = 0$$

and

$$\nabla(V_1 - V_2) = 0$$

Finally, if the gradient of $V_1 - V_2$ is everywhere zero, then $V_1 - V_2$ cannot change with any coordinates and

$$V_1 - V_2 = \text{constant}$$

If we can show that this constant is zero, we shall have accomplished our proof. The constant is easily evaluated by considering a point on the boundary. Here

$V_1 - V_2 = V_{1b} - V_{2b} = 0$, and we see that the constant is indeed zero, and therefore

$$V_1 = V_2$$

giving two identical solutions.

The uniqueness theorem also applies to Poisson's equation, for if $\nabla^2 V_1 = -\rho_v/\epsilon$ and $\nabla^2 V_2 = -\rho_v/\epsilon$, then $\nabla^2(V_1 - V_2) = 0$ as before. Boundary conditions still require that $V_{1b} - V_{2b} = 0$, and the proof is identical from this point.

This constitutes the proof of the uniqueness theorem. Viewed as the answer to a question, "How do two solutions of Laplace's or Poisson's equation compare if they both satisfy the same boundary conditions?" the uniqueness theorem should please us by its ensurance that the answers are identical. Once we can find any method of solving Laplace's or Poisson's equation subject to given boundary conditions, we have solved our problem once and for all. No other method can ever give a different answer.

✔ **D7.2.** Consider the two potential fields $V_1 = y$ and $V_2 = y + e^x \sin y$. (*a*) Is $\nabla^2 V_1 = 0$? (*b*) Is $\nabla^2 V_2 = 0$? (*c*) Is $V_1 = 0$ at $y = 0$? (*d*) Is $V_2 = 0$ at $y = 0$? (*e*) Is $V_1 = \pi$ at $y = \pi$? (*f*) Is $V_2 = \pi$ at $y = \pi$? (*g*) Are V_1 and V_2 identical? (*h*) Why does the uniqueness theorem not apply?

Ans. Yes; yes; yes; yes; yes; yes; no; boundary conditions not given for a *closed* surface

7.3 EXAMPLES OF THE SOLUTION OF LAPLACE'S EQUATION

Several methods have been developed for solving the second-order partial differential equation known as Laplace's equation. The first and simplest method is that of direct integration, and we shall use this technique to work several examples in various coordinate systems in this section. In Sec. 7.5 one other method will be used on a more difficult problem. Additional methods, requiring a more advanced mathematical knowledge, are described in the references given at the end of the chapter.

The method of direct integration is applicable only to problems which are "one-dimensional," or in which the potential field is a function of only one of the three coordinates. Since we are working with only three coordinate systems, it might seem, then, that there are nine problems to be solved, but a little reflection will show that a field which varies only with x is fundamentally the same as a field which varies only with y. Rotating the physical problem a quarter turn is no change. Actually, there are only five problems to be solved, one in cartesian coordinates, two in cylindrical, and two in spherical. We shall enjoy life to the fullest by solving them all.

Example 7.1

Let us assume that V is a function only of x and worry later about which physical problem we are solving when we have a need for boundary conditions. Laplace's equation reduces to

$$\frac{\partial^2 V}{\partial x^2} = 0$$

and the partial derivative may be replaced by an ordinary derivative, since V is not a function of y or z,

$$\frac{d^2 V}{dx^2} = 0$$

We integrate twice, obtaining

$$\frac{dV}{dx} = A$$

and

$$V = Ax + B \tag{12}$$

where A and B are constants of integration. Equation (12) contains two such constants, as we should expect for a second-order differential equation. These constants can be determined only from the boundary conditions.

What boundary conditions should we supply? They are our choice, since no physical problem has yet been specified, with the exception of the original hypothesis that the potential varied only with x. We should now attempt to visualize such a field. Most of us probably already have the answer, but it may be obtained by exact methods.

Since the field varies only with x and is not a function of y and z, then V is a constant if x is a constant or, in other words, the equipotential surfaces are described by setting x constant. These surfaces are parallel planes normal to the x axis. The field is thus that of a parallel-plate capacitor, and as soon as we specify the potential on any two planes, we may evaluate our constants of integration.

To be very general, let $V = V_1$ at $x = x_1$ and $V = V_2$ at $x = x_2$. These values are then substituted into (12), giving

$$V_1 = Ax_1 + B \qquad V_2 = Ax_2 + B$$

$$A = \frac{V_1 - V_2}{x_1 - x_2} \qquad B = \frac{V_2 x_1 - V_1 x_2}{x_1 - x_2}$$

and

$$V = \frac{V_1(x - x_2) - V_2(x - x_1)}{x_1 - x_2}$$

A simpler answer would have been obtained by choosing simpler boundary conditions. If we had fixed $V = 0$ at $x = 0$ and $V = V_0$ at $x = d$, then

$$A = \frac{V_0}{d} \qquad B = 0$$

and

$$\boxed{V = \frac{V_0 x}{d}} \tag{13}$$

Suppose our primary aim is to find the capacitance of a parallel-plate capacitor. We have solved Laplace's equation, obtaining (12) with the two constants A and B. Should they be evaluated or left alone? Presumably we are not interested in the potential field itself, but only in the capacitance, and we may continue successfully with A and B or we may simplify the algebra by a little foresight. Capacitance is given by the ratio of charge to potential difference, so we may choose now the potential *difference* as V_0, which is equivalent to one boundary condition, and then choose whatever second boundary condition seems to help the form of the equation the most. This is the essence of the second set of boundary conditions which produced (13). The potential difference was fixed as V_0 by choosing the potential of one plate zero and the other V_0; the location of these plates was made as simple as possible by letting $V = 0$ at $x = 0$.

Using (13), then, we still need the total charge on either plate before the capacitance can be found. We should remember that when we first solved this capacitor problem in Chap. 5, the sheet of charge provided our starting point. We did not have to work very hard to find the charge, for all the fields were expressed in terms of it. The work then was spent in finding potential difference. Now the problem is reversed (and simplified).

The necessary steps are these, after the choice of boundary conditions has been made:

1. Given V, use $\mathbf{E} = -\nabla V$ to find $\mathbf{E}$.
2. Use $\mathbf{D} = \epsilon \mathbf{E}$ to find $\mathbf{D}$.
3. Evaluate $\mathbf{D}$ at either capacitor plate, $\mathbf{D} = \mathbf{D}_S = D_N \mathbf{a}_N$.
4. Recognize that $\rho_S = D_N$.
5. Find Q by a surface integration over the capacitor plate, $Q = \int_S \rho_S dS$.

Here we have

$$V = V_0 \frac{x}{d}$$

$$\mathbf{E} = -\frac{V_0}{d}\mathbf{a}_x$$

$$\mathbf{D} = -\epsilon \frac{V_0}{d}\mathbf{a}_x$$

$$\mathbf{D}_S = \mathbf{D}\bigg|_{x=0} = -\epsilon\frac{V_0}{d}\mathbf{a}_x$$

$$\mathbf{a}_N = \mathbf{a}_x$$

$$D_N = -\epsilon\frac{V_0}{d} = \rho_S$$

$$Q = \int_S \frac{-\epsilon V_0}{d}\,dS = -\epsilon\frac{V_0 S}{d}$$

and the capacitance is

$$\boxed{C = \frac{|Q|}{V_0} = \frac{\epsilon S}{d}} \tag{14}$$

We shall use this procedure several times in the examples to follow.

Example 7.2

Since no new problems are solved by choosing fields which vary only with y or with z in cartesian coordinates, we pass on to cylindrical coordinates for our next example. Variations with respect to z are again nothing new, and we next assume variation with respect to ρ only. Laplace's equation becomes

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial V}{\partial\rho}\right) = 0$$

or

$$\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{dV}{d\rho}\right) = 0$$

Noting the ρ in the denominator, we exclude $\rho = 0$ from our solution and then multiply by ρ and integrate,

$$\rho\frac{dV}{d\rho} = A$$

rearrange, and integrate again,

$$V = A\ln\rho + B \tag{15}$$

The equipotential surfaces are given by $\rho =$ constant and are cylinders, and the problem is that of the coaxial capacitor or coaxial transmission line. We choose a potential difference of V_0 by letting $V = V_0$ at $\rho = a$, $V = 0$ at $\rho = b$, $b > a$, and obtain

$$\boxed{V = V_0\frac{\ln(b/\rho)}{\ln(b/a)}} \tag{16}$$

from which

$$\mathbf{E} = \frac{V_0}{\rho} \frac{1}{\ln(b/a)} \mathbf{a}_\rho$$

$$D_{N(\rho=a)} = \frac{\epsilon V_0}{a \ln(b/a)}$$

$$Q = \frac{\epsilon V_0 2\pi a L}{a \ln(b/a)}$$

$$\boxed{C = \frac{2\pi\epsilon L}{\ln(b/a)}} \qquad (17)$$

which agrees with our results in Chap. 5.

Example 7.3

Now let us assume that V is a function only of ϕ in cylindrical coordinates. We might look at the physical problem first for a change and see that equipotential surfaces are given by $\phi = \text{constant}$. These are radial planes. Boundary conditions might be $V = 0$ at $\phi = 0$ and $V = V_0$ at $\phi = \alpha$, leading to the physical problem detailed in Fig. 7.1.

Laplace's equation is now

$$\frac{1}{\rho^2} \frac{\partial^2 V}{\partial \phi^2} = 0$$

We exclude $\rho = 0$ and have

$$\frac{d^2 V}{d\phi^2} = 0$$

The solution is

$$V = A\phi + B$$

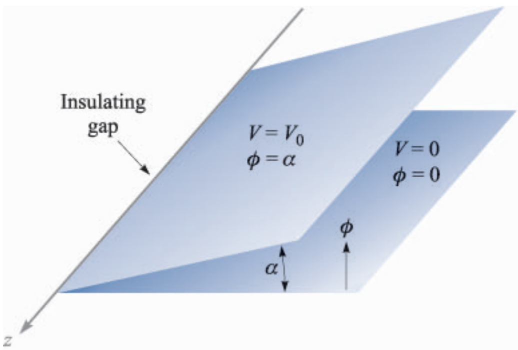


FIGURE 7.1
Two infinite radial planes with an interior angle α. An infinitesimal insulating gap exists at $\rho = 0$. The potential field may be found by applying Laplace's equation in cylindrical coordinates.

The boundary conditions determine A and B, and

$$V = V_0 \frac{\phi}{\alpha} \tag{18}$$

Taking the gradient of (18) produces the electric field intensity,

$$\mathbf{E} = -\frac{V_0 \mathbf{a}_\phi}{\alpha \rho} \tag{19}$$

and it is interesting to note that E is a function of ρ and not of ϕ. This does not contradict our original assumptions, which were restrictions only on the potential field. Note, however, that the *vector* field $\mathbf{E}$ is a function of ϕ.

A problem involving the capacitance of these two radial planes is included at the end of the chapter.

Example 7.4

We now turn to spherical coordinates, dispose immediately of variations with respect to ϕ only as having just been solved, and treat first $V = V(r)$.

The details are left for a problem later, but the final potential field is given by

$$V = V_0 \frac{\frac{1}{r} - \frac{1}{b}}{\frac{1}{a} - \frac{1}{b}} \tag{20}$$

where the boundary conditions are evidently $V = 0$ at $r = b$ and $V = V_0$ at $r = a$, $b > a$. The problem is that of concentric spheres. The capacitance was found previously in Sec. 5.10 (by a somewhat different method) and is

$$C = \frac{4\pi\epsilon}{\frac{1}{a} - \frac{1}{b}} \tag{21}$$

Example 7.5

In spherical coordinates we now restrict the potential function to $V = V(\theta)$, obtaining

$$\frac{1}{r^2 \sin\theta} \frac{d}{d\theta}\left(\sin\theta \frac{dV}{d\theta}\right) = 0$$

We exclude $r = 0$ and $\theta = 0$ or π and have

$$\sin\theta \frac{dV}{d\theta} = A$$

The second integral is then

$$V = \int \frac{A\,d\theta}{\sin\theta} + B$$

which is not as obvious as the previous ones. From integral tables (or a good memory) we have

$$V = A \ln\left(\tan\frac{\theta}{2}\right) + B$$

The equipotential surfaces are cones. Fig. 7.2 illustrates the case where $V = 0$ at $\theta = \pi/2$ and $V = V_0$ at $\theta = \alpha$, $\alpha < \pi/2$. We obtain

$$V = V_0 \frac{\ln\left(\tan\frac{\theta}{2}\right)}{\ln\left(\tan\frac{\alpha}{2}\right)} \tag{22}$$

In order to find the capacitance between a conducting cone with its vertex separated from a conducting plane by an infinitesimal insulating gap and its axis normal to the plane, let us first find the field strength:

$$\mathbf{E} = -\nabla V = -\frac{1}{r}\frac{\partial V}{\partial\theta}\mathbf{a}_\theta = -\frac{V_0}{r\sin\theta\ln\left(\tan\frac{\alpha}{2}\right)}\mathbf{a}_\theta$$

The surface charge density on the cone is then

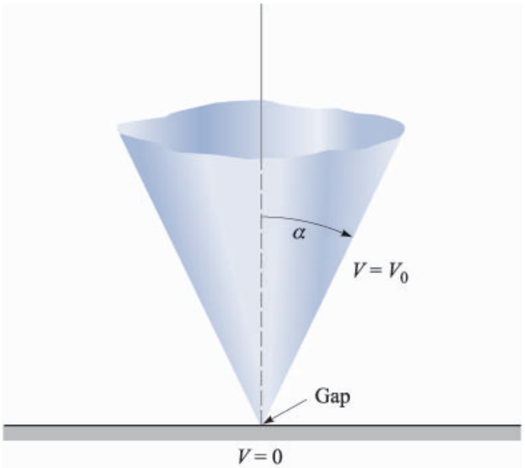


FIGURE 7.2
For the cone $\theta = \alpha$ at V_0 and the plane $\theta = \pi/2$ at $V = 0$, the potential field is given by $V = V_0[\ln(\tan\theta/2)]/[\ln(\tan\alpha/2)]$.

$$\rho_S = \frac{-\epsilon V_0}{r \sin\alpha \ln\left(\tan\frac{\alpha}{2}\right)}$$

producing a total charge Q,

$$Q = \frac{-\epsilon V_0}{\sin\alpha \ln\left(\tan\frac{\alpha}{2}\right)} \int_0^\infty \int_0^{2\pi} \frac{r \sin\alpha \, d\phi \, dr}{r}$$

$$= \frac{-2\pi\epsilon_0 V_0}{\ln\left(\tan\frac{\alpha}{2}\right)} \int_0^\infty dr$$

This leads to an infinite value of charge and capacitance, and it becomes necessary to consider a cone of finite size. Our answer will now be only an approximation, because the theoretical equipotential surface is $\theta = \alpha$, a conical surface extending from $r = 0$ to $r = \infty$, whereas our physical conical surface extends only from $r = 0$ to, say, $r = r_1$. The approximate capacitance is

$$C \doteq \frac{2\pi\epsilon r_1}{\ln\left(\cot\frac{\alpha}{2}\right)} \tag{23}$$

If we desire a more accurate answer, we may make an estimate of the capacitance of the base of the cone to the zero-potential plane and add this amount to our answer above. Fringing, or nonuniform, fields in this region have been neglected and introduce an additional source of error.

✔ **D7.3.** Find $|\mathbf{E}|$ at $P(3, 1, 2)$ for the field of: (*a*) two coaxial conducting cylinders, $V = 50\,\text{V}$ at $\rho = 2\,\text{m}$, and $V = 20\,\text{V}$ at $\rho = 3\,\text{m}$; (*b*) two radial conducting planes, $V = 50\,\text{V}$ at $\phi = 10^\circ$, and $V = 20\,\text{V}$ at $\phi = 30^\circ$.

Ans. 23.4 V/m; 27.2 V/m

7.4 EXAMPLE OF THE SOLUTION OF POISSON'S EQUATION

To select a reasonably simple problem which might illustrate the application of Poisson's equation, we must assume that the volume charge density is specified. This is not usually the case, however; in fact, it is often the quantity about which we are seeking further information. The type of problem which we might encounter later would begin with a knowledge only of the boundary values of the potential, the electric field intensity, and the current density. From these we would have to apply Poisson's equation, the continuity equation, and some relationship expressing the forces on the charged particles, such as the Lorentz force equation or the diffusion equation, and solve the whole system of equations

simultaneously. Such an ordeal is beyond the scope of this text, and we shall therefore assume a reasonably large amount of information.

As an example, let us select a *pn* junction between two halves of a semiconductor bar extending in the x direction. We shall assume that the region for $x < 0$ is doped p type and that the region for $x > 0$ is n type. The degree of doping is identical on each side of the junction. To review qualitatively some of the facts about the semiconductor junction, we note that initially there are excess holes to the left of the junction and excess electrons to the right. Each diffuses across the junction until an electric field is built up in such a direction that the diffusion current drops to zero. Thus, to prevent more holes from moving to the right, the electric field in the neighborhood of the junction must be directed to the left; E_x is negative there. This field must be produced by a net positive charge to the right of the junction and a net negative charge to the left. Note that the layer of positive charge consists of two parts—the holes which have crossed the junction and the positive donor ions from which the electrons have departed. The negative layer of charge is constituted in the opposite manner by electrons and negative acceptor ions.

The type of charge distribution which results is shown in Fig. 7.3*a*, and the negative field which it produces is shown in Fig. 7.3*b*. After looking at these two figures, one might profitably read the previous paragraph again.

A charge distribution of this form may be approximated by many different expressions. One of the simpler expressions is

$$\rho_v = 2\rho_{v0} \operatorname{sech} \frac{x}{a} \tanh \frac{x}{a} \tag{24}$$

which has a maximum charge density $\rho_{v,max} = \rho_{v0}$ that occurs at $x = 0.881a$. The maximum charge density ρ_{v0} is related to the acceptor and donor concentrations N_a and N_d by noting that all the donor and acceptor ions in this region (the *depletion* layer) have been stripped of an electron or a hole, and thus

$$\rho_{v0} = eN_a = eN_d$$

Let us now solve Poisson's equation,

$$\nabla^2 V = -\frac{\rho_v}{\epsilon}$$

subject to the charge distribution assumed above,

$$\frac{d^2V}{dx^2} = -\frac{2\rho_{v0}}{\epsilon} \operatorname{sech} \frac{x}{a} \tanh \frac{x}{a}$$

in this one-dimensional problem in which variations with y and z are not present. We integrate once,

$$\frac{dV}{dx} = \frac{2\rho_{v0}a}{\epsilon} \operatorname{sech} \frac{x}{a} + C_1$$

and obtain the electric field intensity,

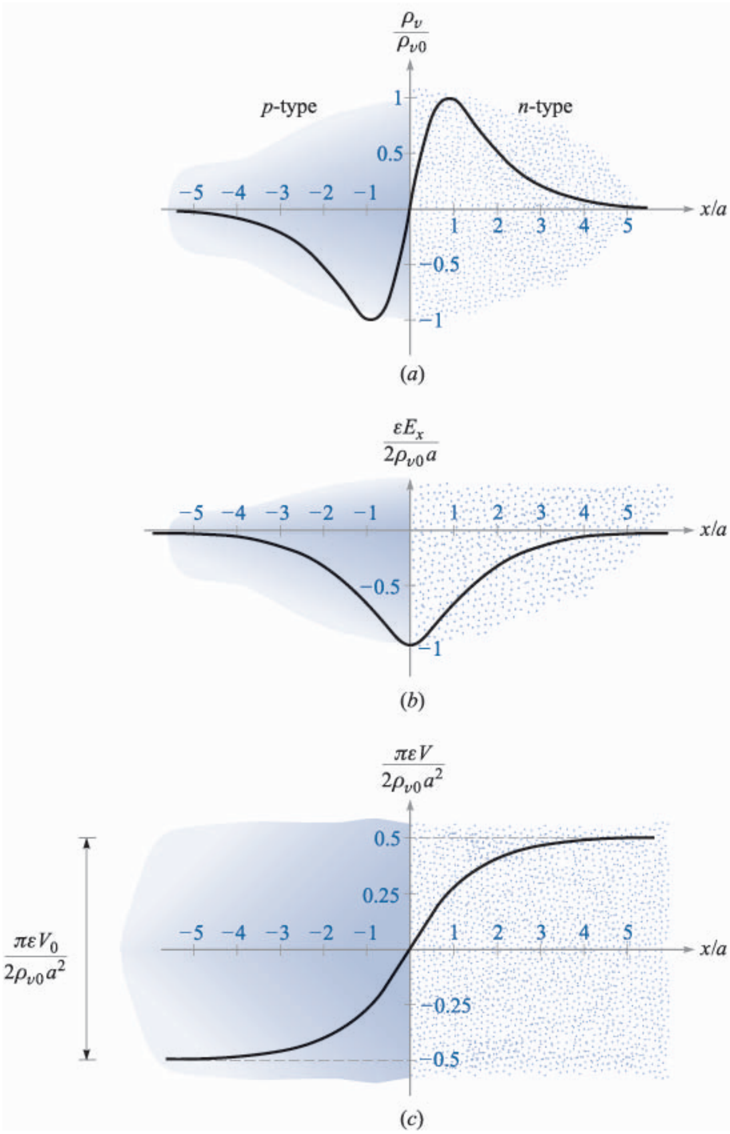


FIGURE 7.3
(*a*) The charge density, (*b*) the electric field intensity, and (*c*) the potential are plotted for a pn junction as functions of distance from the center of the junction. The *p*-type material is on the left, and the *n*-type is on the right.

$$E_x = -\frac{2\rho_{v0}a}{\epsilon}\operatorname{sech}\frac{x}{a} - C_1$$

To evaluate the constant of integration C_1, we note that no net charge density and no fields can exist *far* from the junction. Thus, as $x \to \pm\infty$, E_x must approach zero. Therefore $C_1 = 0$, and

$$E_x = -\frac{2\rho_{v0}a}{\epsilon}\operatorname{sech}\frac{x}{a} \tag{25}$$

Integrating again,

$$V = \frac{4\rho_{v0}a^2}{\epsilon}\tan^{-1} e^{x/a} + C_2$$

Let us arbitrarily select our zero reference of potential at the center of the junction, $x = 0$,

$$0 = \frac{4\rho_{v0}a^2}{\epsilon}\frac{\pi}{4} + C_2$$

and finally,

$$V = \frac{4\rho_{v0}a^2}{\epsilon}\left(\tan^{-1} e^{x/a} - \frac{\pi}{4}\right) \tag{26}$$

Fig. 7.3 shows the charge distribution (*a*), electric field intensity (*b*), and the potential (*c*), as given by (24), (25), and (26), respectively.

The potential is constant once we are a distance of about $4a$ or $5a$ from the junction. The total potential difference V_0 across the junction is obtained from (26),

$$V_0 = V_{x\to\infty} - V_{x\to-\infty} = \frac{2\pi\rho_{v0}a^2}{\epsilon} \tag{27}$$

This expression suggests the possibility of determining the total charge on one side of the junction and then using (27) to find a junction capacitance. The total positive charge is

$$Q = S\int_0^\infty 2\rho_{v0}\operatorname{sech}\frac{x}{a}\tanh\frac{x}{a}\,dx = 2\rho_{v0}aS$$

where S is the area of the junction cross section. If we make use of (27) to eliminate the distance parameter a, the charge becomes

$$Q = S\sqrt{\frac{2\rho_{v0}\epsilon V_0}{\pi}} \tag{28}$$

Since the total charge is a function of the potential difference, we have to be careful in defining a capacitance. Thinking in "circuit" terms for a moment,

$$I = \frac{dQ}{dt} = C\frac{dV_0}{dt}$$

and thus

$$C = \frac{dQ}{dV_0}$$

By differentiating (28) we therefore have the capacitance,

$$C = \sqrt{\frac{\rho_{v0}\epsilon}{2\pi V_0}} S = \frac{\epsilon S}{2\pi a} \tag{29}$$

The first form of (29) shows that the capacitance varies inversely as the square root of the voltage. That is, a higher voltage causes a greater separation of the charge layers and a smaller capacitance. The second form is interesting in that it indicates that we may think of the junction as a parallel-plate capacitor with a "plate" separation of $2\pi a$. In view of the dimensions of the region in which the charge is concentrated, this is a logical result.

Poisson's equation enters into any problem involving volume charge density. Besides semiconductor diode and transistor models, we find that vacuum tubes, magnetohydrodynamic energy conversion, and ion propulsion require its use in constructing satisfactory theories.

✔ **D7.4.** In the neighborhood of a certain semiconductor junction the volume charge density is given by $\rho_v = 750 \operatorname{sech} 10^6 \pi x \tanh \pi x \ \mathrm{C/m^3}$. The dielectric constant of the semiconductor material is 10 and the junction area is $2 \times 10^{-7}\ \mathrm{m^2}$. Find: (*a*) V_0; (*b*) C; (*c*) E at the junction.

Ans. 2.70 V; 8.85 pF; 2.70 MV/m

✔ **D7.5.** Given the volume charge density $\rho_v = -2 \times 10^7 \epsilon_0 \sqrt{x}\ \mathrm{C/m^3}$ in free space, let $V = 0$ at $x = 0$ and $V = 2$ V at $x = 2.5$ mm. At $x = 1$ mm, find: (*a*) V; (*b*) E_x.

Ans. 0.302 V; −555 V/m

7.5 PRODUCT SOLUTION OF LAPLACE'S EQUATION

In this section we are confronted with the class of potential fields which vary with more than one of the three coordinates. Although our examples are taken in the cartesian coordinate system, the general method is applicable to the other coordinate systems. We shall avoid those applications, however, because the potential fields are given in terms of more advanced mathematical functions, such as Bessel functions and spherical and cylindrical harmonics, and our interest now does not lie with new mathematical functions but with the techniques and methods of solving electrostatic field problems.

We may give ourselves a general class of problems by specifying merely that the potential is a function of x and y alone, so that

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0 \tag{30}$$

We now assume that the potential is expressible as the *product* of a function of x alone and a function of y alone. It might seem that this prohibits too many solutions, such as $V = x + y$, or any sum of a function of x and a function of y, but we should realize that Laplace's equation is linear and the sum of any two solutions is also a solution. We could treat $V = x + y$ as the sum of $V_1 = x$ and $V_2 = y$, where each of these latter potentials is now a (trivial) product solution.

Representing the function of x by X and the function of y by Y, we have

$$V = X\,Y \tag{31}$$

which is substituted into (30),

$$Y\frac{\partial^2 X}{\partial x^2} + X\frac{\partial^2 Y}{\partial y^2} = 0$$

Since X does not involve y and Y does not involve x, ordinary derivatives may be used,

$$Y\frac{d^2 X}{dx^2} + X\frac{d^2 Y}{dy^2} = 0 \tag{32}$$

Equation (32) may be solved by separating the variables through division by XY, giving

$$\frac{1}{X}\frac{d^2 X}{dx^2} + \frac{1}{Y}\frac{d^2 Y}{dy^2} = 0$$

or

$$\frac{1}{X}\frac{d^2 X}{dx^2} = -\frac{1}{Y}\frac{d^2 Y}{dy^2}$$

Now we need one of the cleverest arguments of mathematics: since $(1/X)d^2X/dx^2$ involves no y and $-(1/Y)d^2Y/dy^2$ involves no x, and since the two quantities are equal, then $(1/X)d^2X/dx^2$ cannot be a function of x either, and similarly, $-(1/Y)d^2Y/dy^2$ cannot be a function of y! In other words, we have shown that each of these terms must be a constant. For convenience, let us call this constant α^2,

$$\frac{1}{X}\frac{d^2 X}{dx^2} = \alpha^2 \tag{33}$$

$$-\frac{1}{Y}\frac{d^2 Y}{dy^2} = \alpha^2 \tag{34}$$

The constant α^2 is called the *separation constant*, because its use results in separating one equation into two simpler equations.

Equation (33) may be written as

$$\frac{d^2X}{dx^2} = \alpha^2 X \tag{35}$$

and must now be solved. There are several methods by which a solution may be obtained. The first method is experience, or recognition, which becomes more powerful with practice. We are just beginning and can barely recognize Laplace's equation itself. The second method might be that of direct integration, when applicable, of course. Applying it here, we should write

$$d\left(\frac{dX}{dx}\right) = \alpha^2 X\, dx$$

$$\frac{dX}{dx} = \alpha^2 \int X\, dx$$

and then pass on to the next method, for X is some unknown function of x, and the method of integration is not applicable here. The third method we might describe as intuition, common sense, or inspection. It involves taking a good look at the equation, perhaps putting the operation into words. This method will work on (35) for some of us if we ask ourselves, "What function has a second derivative which has the same form as the function itself, except for multiplication by a constant?" The answer is the exponential function, of course, and we could go on from here to construct the solution. Instead, let us work with those of us whose intuition is suffering from exposure and apply a very powerful but long method, the infinite-power-series substitution.

We assume hopefully that X may be represented by

$$X = \sum_{a=0}^{\infty} a_n x^n$$

and substitute into (35), giving

$$\frac{d^2X}{dx^2} = \sum_0^{\infty} n(n-1)a_n x^{n-2} = \alpha^2 \sum_0^{\infty} a_n x^n$$

If these two different infinite series are to be equal for all x, they must be identical, and the coefficients of like powers of x may be equated term by term. Thus

$$2 \times 1 \times a_2 = \alpha^2 a_0$$

$$3 \times 2 \times a_3 = \alpha^2 a_1$$

and in general we have the recurrence relationship

$$(n+2)(n+1)a_{n+2} = \alpha^2 a_n$$

The even coefficients may be expressed in terms of a_0 as

$$a_2 = \frac{\alpha^2}{1 \times 2} a_0$$

$$a_4 = \frac{\alpha^2}{3 \times 4} a_2 = \frac{\alpha^4}{4!} a_0$$

$$a_6 = \frac{\alpha^6}{6!} a_0$$

and, in general, for n even, as

$$a_n = \frac{\alpha^n}{n!} a_0 \quad (n \text{ even})$$

For odd values of n, we have

$$a_3 = \frac{\alpha^2}{2 \times 3} a_1 = \frac{\alpha^3}{3!} \frac{a_1}{\alpha}$$

$$a_5 = \frac{\alpha^5}{5!} \frac{a_1}{\alpha}$$

and in general, for n odd,

$$a_n = \frac{\alpha^n}{n!} \frac{a_1}{\alpha} \quad (n \text{ odd})$$

Substituting back into the original power series for X, we obtain

$$X = a_0 \sum_{0,\text{even}}^{\infty} \frac{\alpha^n}{n!} x^n + \frac{a_1}{\alpha} \sum_{1,\text{odd}}^{\infty} \frac{\alpha^n}{n!} x^n$$

or

$$X = a_0 \sum_{0,\text{even}}^{\infty} \frac{(\alpha x)^n}{n!} + \frac{a_1}{\alpha} \sum_{1,\text{odd}}^{\infty} \frac{(\alpha x)^n}{n!}$$

Although the sum of these two infinite series is the solution of the differential equation in x, the form of the solution may be improved immeasurably by recognizing the first series as the hyperbolic cosine,

$$\cosh \alpha x = \sum_{0,\text{even}}^{\infty} \frac{(\alpha x)^n}{n!} = 1 + \frac{(\alpha x)^2}{2!} + \frac{(\alpha x)^4}{4!} + \ldots$$

and the second series as the hyperbolic sine,

$$\sinh \alpha x = \sum_{1,\text{odd}}^{\infty} \frac{(\alpha x)^n}{n!} = \alpha x + \frac{(\alpha x)^3}{3!} + \frac{(\alpha x)^5}{5!} + .$$

The solution may therefore be written as

$$X = a_0 \cosh \alpha x + \frac{a_1}{\alpha} \sinh \alpha x$$

or

$$X = A \cosh \alpha x + B \sinh \alpha x$$

where the slightly simpler terms A and B have replaced a_0 and a_1/α, respectively, and are the two constants which must be evaluated in terms of the boundary conditions. The separation constant is not an arbitrary constant as far as the solution of (35) is concerned, for it appears in that equation.

An alternate form of the solution is obtained by expressing the hyperbolic functions in terms of exponentials, collecting terms, and selecting new arbitrary constants, A' and B',

$$X = A'e^{\alpha x} + B'e^{-\alpha x}$$

Turning our attention now to (34), we see the solution proceeds along similar lines, leading to two power series representing the sine and cosine, and we have

$$Y = C \cos \alpha y + D \sin \alpha y$$

from which the potential is

$$V = X\,Y = (A \cosh \alpha x + B \sinh \alpha x)(C \cos \alpha y + D \sin \alpha y) \tag{36}$$

Before describing a physical problem and forcing the constants appearing in (36) to fit the boundary conditions prescribed, let us consider the physical nature of the potential field given by a simple choice of these constants. Letting $A = 0$, $C = 0$, and $BD = V_1$, we have

$$V = V_1 \sinh \alpha x \sin \alpha y \tag{37}$$

The sinh αx factor is zero at $x = 0$ and increases smoothly with x, soon becoming nearly exponential in form, since

$$\sinh \alpha x = \tfrac{1}{2}(e^{\alpha x} - e^{-\alpha x})$$

The sin αy term causes the potential to be zero at $y = 0$, $y = \pi/\alpha$, $y = 2\pi/\alpha$, and so forth. We therefore may place zero-potential conducting planes at $x = 0$, $y = 0$, and $y = \pi/\alpha$. Finally, we can describe the V_1 equipotential surface by setting $V = V_1$ in (37), obtaining

$$\sinh \alpha x \sin \alpha y = 1$$

or

$$\alpha y = \sin^{-1} \frac{1}{\sinh \alpha x}$$

This is not a familiar equation, but a hand calculator or a set of tables can furnish enough material values to allow us to plot αy as a function of αx. Such a

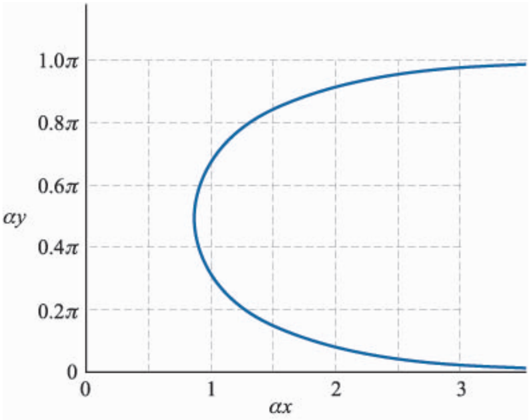


FIGURE 7.4
A graph of the double-valued function $\alpha y = \sin^{-1}(1/\sinh \alpha x)$, $0 < \alpha y < \pi$.

curve is shown in Fig. 7.4. Note that the curve is double-valued and symmetrical about the line $\alpha y = \pi/2$ when αy is restricted to the interval between 0 and π. The information of Fig. 7.4 is transferred directly to the $V = 0$ and $V = V_1$ equipotential conducting surfaces in Fig. 7.5. The surfaces are shown in cross section, since the potential is not a function of z.

It is very unlikely that we shall ever be asked to find the potential field of these peculiarly shaped electrodes, but we should bear in mind the possibility of combining a number of the fields having the form given by (36) or (37) and thus satisfying the boundary conditions of a more practical problem. We close this chapter with such an example.

The problem to be solved is that shown in Fig. 7.6. The boundary conditions shown are $V = 0$ at $x = 0$, $y = 0$, and $y = b$, and $V = V_0$ at $x = d$ for all y

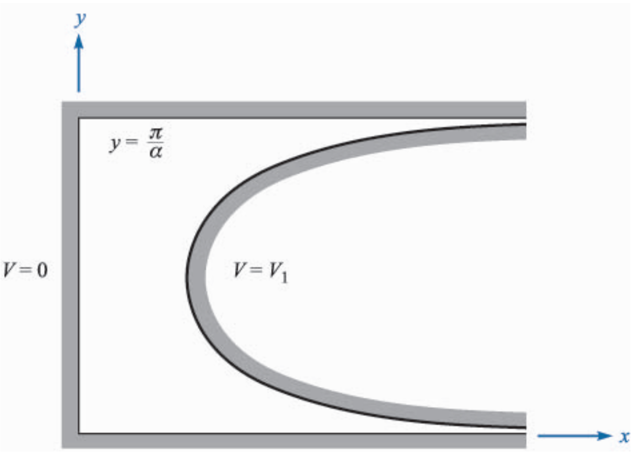


FIGURE 7.5
Cross section of the $V = 0$ and $V = V_1$ equipotential surfaces for the potential field $V = V_1 \sinh \alpha x \sin \alpha y$.

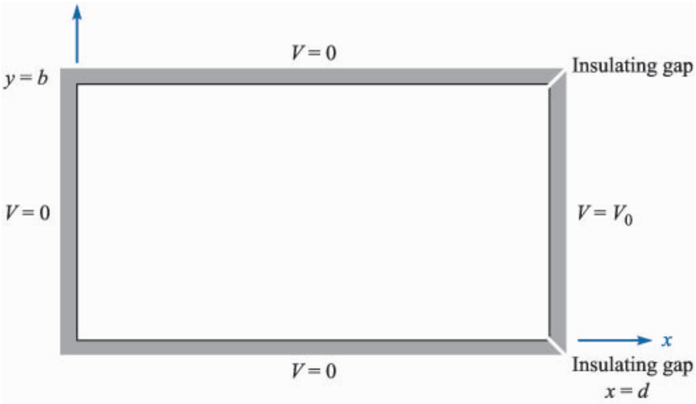


FIGURE 7.6
Potential problem requiring an infinite summation of fields of the form $V = V_1 \sinh \alpha x \sin \alpha y$. A similar configuration was analyzed by the iteration method in Chap. 6.

between 0 and b. It is immediately apparent that the potential field given by(37) and outlined in Fig. 7.5 satisfies two of the four boundary conditions. A third condition, $V = 0$ at $y = b$, may be satisfied by the choice of a, for the substitution of these values of (37) leads to the equation

$$0 = V_1 \sinh \alpha x \sin \alpha b$$

which may be satisfied by setting

$$\alpha b = m\pi \qquad (m = 1, 2, 3, \ldots)$$

or

$$\alpha = \frac{m\pi}{b}$$

The potential function

$$V = V_1 \sinh \frac{m\pi x}{b} \sin \frac{m\pi y}{b} \tag{38}$$

thus produces the correct potential at $x = 0$, $y = 0$, and $y = b$, regardless of the choice of m or the value of V_1. It is impossible to choose m or V_1 in such a way that $V = V_0$ at $x = d$ for each and every value of y between 0 and b. We must combine an infinite number of these fields, each with a different value of m and a corresponding value of V_1,

$$V = \sum_{m=0}^{\infty} V_{1m} \sinh \frac{m\pi x}{b} \sin \frac{m\pi y}{b}$$

The subscript on V_{1m} indicates that this amplitude factor will have a different value for each different value of m. Applying the last boundary condition now,

$$V_0 = \sum_{m=0}^{\infty} V_{1m} \sinh \frac{m\pi d}{d} \sin \frac{m\pi y}{b} \qquad (0 < y < b, m = 1, 2, \ldots)$$

Since $V_{1m} \sinh (m\pi d/b)$ is a function only of m, we may simplify the expression by replacing this factor by c_m,

$$V_0 = \sum_{m=0}^{\infty} c_m \sin \frac{m\pi y}{b} \qquad (0 < y < b, m = 1, 2, \ldots)$$

This is a Fourier sine series, and the c_m coefficients may be determined by the standard Fourier-series methods[1] if we can interpret V_0 as a periodic function of y. Since our physical problem is bounded by conducting planes at $y = 0$ and $y = b$, and our interest in the potential does not extend outside of this region, we may *define* the potential at $x = d$ for y *outside* of the range 0 to b in any manner we choose. Probably the simplest periodic expression is obtained by selecting the interval $0 < y < b$ as the half-period and choosing $V = -V_0$ in the adjacent half-period, or

$$V = V_0 \qquad (x = d, 0 < y < b)$$
$$V = -V_0 \qquad (x = d, b < y < 2b)$$

The c_m coefficients are then

$$c_m = \frac{1}{b}\left[\int_0^b V_0 \sin \frac{m\pi y}{b}\, dy + \int_b^{2b} (-V_0) \sin \frac{m\pi y}{b}\, dy\right]$$

leading to

$$c_m = \frac{4V_0}{m\pi} \qquad (m \text{ odd})$$
$$= 0 \qquad (m \text{ even})$$

However, $c_m = V_{1m} \sinh (m\pi d/b)$, and therefore

$$V_{1m} = \frac{4V_0}{m\pi \sinh (m\pi d/b)} \qquad (m \text{ odd only})$$

which may be substituted into (38) to give the desired potential function,

$$V = \frac{4V_0}{\pi} \sum_{1,\text{odd}}^{\infty} \frac{1}{m} \frac{\sinh (m\pi x/b)}{\sinh (m\pi d/b)} \sin \frac{m\pi y}{b} \tag{39}$$

The map of this field may be obtained by evaluating (39) at a number of points and drawing equipotentials by interpolation between these points. If we let $b = d$ and $V_0 = 100$, the problem is identical with that used as the example in

[1] Fourier series are discussed in almost every electrical engineering text on circuit theory. The authors are partial to the Hayt and Kemmerly reference given in the Suggested References at the end of the chapter.

the discussion of the iteration method. Checking one of the grid points in that problem, we let $x = d/4 = b/4$, $y = b/2 = d/2$, and $V_0 = 100$ and obtain

$$V = \frac{400}{\pi} \sum_{1,\text{odd}}^{\infty} \frac{1}{m} \frac{\sinh(m\pi/4)}{\sinh m\pi} \sin \frac{m\pi}{2}$$

$$= \frac{400}{\pi}\left(\frac{\sinh(\pi/4)}{\sinh \pi} - \frac{1}{3}\frac{\sinh(3\pi/4)}{\sinh 3\pi} + \frac{1}{5}\frac{\sinh(5\pi/4)}{\sinh 5\pi} - \dots\right)$$

$$= \frac{400}{\pi}\left(\frac{0.8687}{11.549} - \frac{5.228}{3 \times 6195.8} + \dots\right)$$

$$= 9.577 - 0.036 + \dots$$

$$= 9.541 \text{ V}$$

The equipotentials are drawn for increments of 10 V in Fig. 7.7, and flux lines have been added graphically to produce a curvilinear map.

The material covered in this discussion of the product solution was more difficult than much of the preceding work, and moreover, it presented three new ideas. The first new technique was the assumption that the potential might be expressed as the product of a function of x and a function of y, and the resultant separation of Laplace's equation into two simpler ordinary differential equations. The second new approach was employed when an infinite-power-series solution was assumed as the solution for one of the ordinary differential equations. Finally, we considered an example which required the combination of an infinite number of simpler product solutions, each having a different amplitude and a different variation in one of the coordinate directions. All these techniques

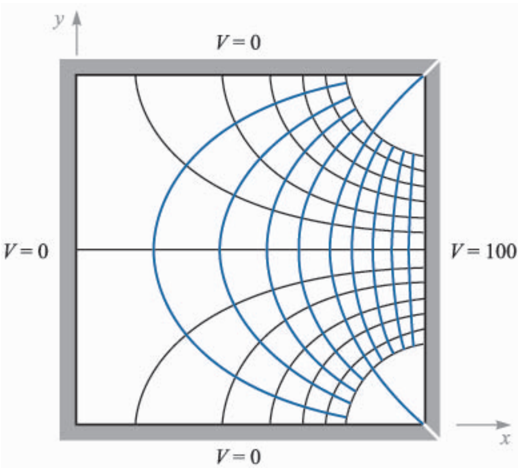


FIGURE 7.7
The field map corresponding to $V = \frac{4V_0}{\pi} \sum_{1,\text{odd}}^{\infty} \frac{1}{m} \frac{\sinh(m\pi x/b)}{\sinh(m\pi d/b)} \sin \frac{m\pi y}{b}$ with $b = d$ and $V_0 = 100$ V.

are very powerful. They are useful in all coordinate systems, and they can be used in problems in which the potential varies with all three coordinates.

We have merely introduced the subject here, and more information can be obtained from the references below, several of which devote hundreds of pages to the solution of Laplace's equation.

SUGGESTED REFERENCES

1. Dekker, A. J.: (see Suggested References for Chap. 5).
2. Hayt, W. H., Jr., and J. E. Kemmerly: "Engineering Circuit Analysis," 5th ed., McGraw-Hill Book Company, New York, 1993.
3. Push, E. M., and E. W. Pugh: "Principles of Electricity and Magnetism," 2d ed., Addison-Wesley Publishing Co., Reading, Mass., 1970. This text provides the physicist's view of electricity and magnetism, but electrical engineering students should find it easy to read. The solution to Laplace's equation by a number of methods is discussed in chap. 4.
4. Ramo, S., J. R. Whinnery, and T. Van Duzer: (see Suggested References for Chap. 6). A more complete and advanced discussion of methods of solving Laplace's equation is given in chap. 7.
5. Seeley, S., and A. D. Poularikas: "Electromagnetics: Classical and Modern Theory and Applications," Marcel Dekker, Inc., New York, 1979. Several examples of the solution of Laplace's equation by separation of variables appear in chap. 4.
6. Smythe, W. R.: "Static and Dynamic Electricity," 3d ed., McGraw-Hill Book Company, New York, 1968. An advanced treatment of potential theory is given in chap. 4.
7. Weber, E.: (see Suggested References for Chap. 6). There are a tremendous number of potential solutions given with the original references.

PROBLEMS

7.1 Let $V = 2xy^2z^3$ and $\epsilon = \epsilon_0$. Given point $P(1, 2, -1)$, find: (*a*) V at P; (*b*) $\mathbf{E}$ at P; (*c*) ρ_v at P; (*d*) the equation of the equipotential surface passing through P; (*e*) the equation of the streamline passing through P. (*f*) Does V satisfy Laplace's equation?

7.2 A potential field V exists in a region where $\epsilon = f(x)$. Find $\nabla^2 V$ if $\rho_v = 0$.

7.3 Let $V(x, y) = 4e^{2x} + f(x) - 3y^2$ in a region of free space where $\rho_v = 0$. It is known that both E_x and V are zero at the origin. Find $f(x)$ and $V(x, y)$.

7.4 Given the potential field $V = A \ln\left(\tan^2 \dfrac{\theta}{2}\right) + B$: (*a*) show that $\nabla^2 V = 0$; (*b*) select A and B so that $V = 100\,\text{V}$ and $E_\theta = 500\,\text{V/m}$ at $P(r = 5, \theta = 60^\circ, \phi = 45^\circ)$.

7.5 Given the potential field $V = (A\rho^4 + B\rho^{-4}) \sin 4\phi$: (*a*) show that $\nabla^2 V = 0$; (*b*) select A and B so that $V = 100\,\text{V}$ and $|\mathbf{E}| = 500\,\text{V/m}$ at $P(\rho = 1, \phi = 22.5^\circ, z = 2)$.

7.6 If $V = \dfrac{20 \sin\theta}{r^3}$ V in free space, find: (*a*) ρ_v at $P(r = 2, \theta = 30^\circ, \phi = 0)$; (*b*) the total charge within the spherical shell $1 < r < 2\,\text{m}$.

7.7 Let $V = \dfrac{\cos 2\phi}{\rho}$ V in free space. (*a*) Find the volume charge density at point $A(\frac{1}{2}, 60^\circ, 1)$. (*b*) Find the surface charge density on a conductor surface passing through the point $B(2, 30^\circ, 1)$.

7.8 Let $V_1(r, \theta, \phi) = \dfrac{20}{r}$ and $V_2(r, \theta, \phi) = \dfrac{4}{r} + 4$. (*a*) State whether V_1 and V_2 satisfy Laplace's equation. (*b*) Evaluate V_1 and V_2 on the closed surface $r = 4$. (*c*) Conciliate your results with the uniqueness theorem.

7.9 The functions $V_1(\rho, \phi, z)$ and $V_2(\rho, \phi, z)$ both satisfy Laplace's equation in the region $a < \rho < b$, $0 \le \phi < 2\pi$, $-L < z < L$; each is zero on the surfaces $\rho = b$ for $-L < z < L$; $z = -L$ for $a < \rho < b$; and $z = L$ for $a < \rho < b$; and each is 100 V on the surface $\rho = a$ for $-L < z < L$. (*a*) In the region specified above, is Laplace's equation satisfied by the functions $V_1 + V_2$, $V_1 - V_2$, $V_1 + 3$, and $V_1 V_2$? (*b*) On the boundary surfaces specified, are the potential values given above obtained from the functions $V_1 + V_2$, $V_1 - V_2$, $V_1 + 3$, and $V_1 V_2$? (*c*) Are the functions V_2, $V_1 + V_2$, $V_1 + 3$, and $V_1 V_2$ identical with V_1?

7.10 Conducting planes at $z = 2\,\text{cm}$ and $z = 8\,\text{cm}$ are held at potentials of -3 V and 9 V, respectively. The region between the plates is filled with a perfect dielectric with $\epsilon = 5\epsilon_0$. Find and sketch: (*a*) $V(z)$; (*b*) $E_z(z)$; (*c*) $D_z(z)$.

7.11 The conducting planes $2x + 3y = 12$ and $2x + 3y = 18$ are at potentials of 100 V and 0, respectively. Let $\epsilon = \epsilon_0$ and find: (*a*) V at $P(5, 2, 6)$; (*b*) $\mathbf{E}$ at P.

7.12 Conducting cylinders at $\rho = 2\,\text{cm}$ and $\rho = 8\,\text{cm}$ in free space are held at potentials of 60 mV and -30 mV, respectively. (*a*) Find $V(\rho)$. (*b*) Find $E_\rho(\rho)$. (*c*) Find the surface on which $V = 30\,\text{mV}$.

7.13 Coaxial conducting cylinders are located at $\rho = 0.5\,\text{cm}$ and $\rho = 1.2\,\text{cm}$. The region between the cylinders is filled with a homogeneous perfect dielectric. If the inner cylinder is at 100 V and the outer at 0 V, find: (*a*) the location of the 20-V equipotential surface; (*b*) $E_{\rho\,\text{max}}$; (*c*) ϵ_R if the charge per meter length on the inner cylinder is 20 nC/m.

7.14 Two semi-infinite planes are located at $\phi = -\alpha$ and $\phi = \alpha$, where $\alpha < \pi/2$. A narrow insulating strip separates them along the z axis. The potential at $\phi = -\alpha$ is V_0, while $V = 0$ at $\phi = \alpha$. (*a*) Find $V(\phi)$ in terms of α and V_0. (*b*) Find E_ϕ at $\phi = 20^\circ$, $\rho = 2\,\text{cm}$, if $V_0 = 100\,V$ and $\alpha = 30^\circ$.

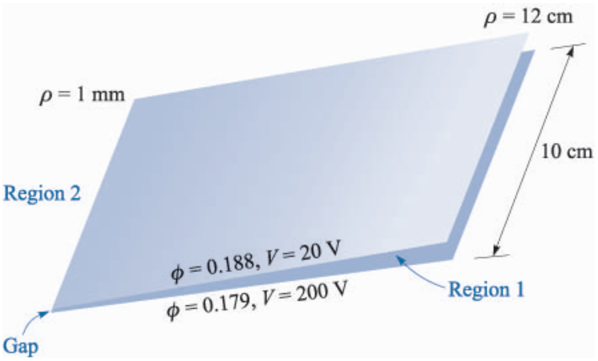


FIGURE 7.8
See Prob. 15.

7.15 The two conducting planes illustrated in Fig. 7.8 are defined by $0.001 < \rho < 0.120$ m, $0 < z < 0.1$ m, $\phi = 0.179$ and 0.188 rad. The medium surrounding the planes is air. For region 1, $0.179 < \phi < 0.188$, neglect fringing and find: (*a*) $V(\phi)$; (*b*) $\mathbf{E}(\rho)$; (*c*) $\mathbf{D}(\rho)$; (*d*) ρ_S on the upper surface of the lower plane; (*e*) Q on the upper surface of the lower plane. (*f*) Repeat (*a*) to (*c*) for region 2 by letting the location of the upper plane be $\phi = 0.188 - 2\pi$, and then find ρ_S and Q on the lower surface of the lower plane. (*g*) Find the total charge on the lower plane and the capacitance between the planes.

7.16 (*a*) Solve Laplace's equation for the potential field in the homogeneous region between two concentric conducting spheres with radii a and b, $b > a$, if $V = 0$ at $r = b$, and $V = V_0$ at $r = a$. (*b*) Find the capacitance between them.

7.17 Concentric conducting spheres are located at $r = 5$ mm and $r = 20$ mm. The region between the spheres is filled with a perfect dielectric. If the inner sphere is at 100 V and the outer at 0 V: (*a*) find the location of the 20-V equipotential surface; (*b*) find $E_{r,\max}$; (*c*) find ϵ_R if the surface charge density on the inner sphere is $100\,\mu\text{C/m}^2$.

7.18 Concentric conducting spheres have radii of 1 and 5 cm. There is a perfect dielectric for which $\epsilon_R = 3$ between them. The potential of the inner sphere is 2 V and that of the outer is -2 V. Find: (*a*) $V(r)$; (*b*) $\mathbf{E}(r)$; (*c*) V at $r = 3$ cm; (*d*) the location of the 0-V equipotential surface; (*e*) the capacitance between the spheres.

7.19 Two coaxial conducting cones have their vertices at the origin and the z axis as their axis. Cone A has the point $A(1, 0, 2)$ on its surface, while cone B has the point $B(0, 3, 2)$ on its surface. Let $V_A = 100$ V and $V_B = 20$ V. Find: (*a*) α for each cone; (*b*) V at $P(1, 1, 1)$.

7.20 A potential field in free space is given as $V = 100 \ln[\tan(\theta/2)] + 50$ V. (*a*) Find the maximum value of $|\mathbf{E}_\theta|$ on the surface $\theta = 40^\circ$ for $0.1 < r < 0.8$ m, $60^\circ < \phi < 90^\circ$. (*b*) Describe the surface $V = 80$ V.

7.21 In free space, let $\rho_v = 200\epsilon_0/r^{2.4}$. (*a*) Use Poisson's equation to find $V(r)$ if it is assumed that $r^2E_r \to 0$ when $r \to 0$, and also that $V \to 0$ as $r \to \infty$. (*b*) Now find $V(r)$ by using Gauss's law and a line integral.

7.22 Let the volume charge density in Fig. 7.3*a* be given by $\rho_v = \rho_{v0}(x/a)e^{-|x|a}$. (*a*) Determine $\rho_{v,\max}$ and $\rho_{v,\min}$ and their locations. (*b*) Find E_x and $V(x)$ if $V(0) = 0$ and $E_x \to 0$ as $x \to \infty$. (*c*) Use a development similar to that of Sec. 7.4 to show that $C = dQ/dV_0 = \epsilon_0 S/(4\sqrt{2}a)$.

7.23 A rectangular trough is formed by four conducting planes located at $x = 0$ and 8 cm and $y = 0$ and 5 cm in air. The surface at $y = 5$ cm is at a potential of 100 V, the other three are at zero potential, and the necessary gaps are placed at two corners. Find the potential at $x = 3$ cm, $y = 4$ cm.

7.24 The four sides of a square trough are held at potentials of 0, 20, −30, and 60 V; the highest and lowest potentials are on opposite sides. Find the potential at the center of the trough.

7.25 In Fig. 7.7 change the right side so that the potential varies linearly from 0 at the bottom of that side to 100 V at the top. Solve for the potential at the center of the trough.

7.26 If X is a function of x and $X'' + (x-1)X' - 2X = 0$, assume a solution in the form of an infinite power series and determine numerical values for a_2 to a_8 if $a_0 = 1$ and $a_1 = -1$.

7.27 It is known that $V = XY$ is a solution of Laplace's equation, where X is a function of x alone and Y is a function of y alone. Determine which of the following potential functions are also solutions of Laplace's equation: (*a*) $V = 100X$; (*b*) $V = 50XY$; (*c*) $V = 2XY + x - 3y$; (*d*) $V = xXY$; (*e*) $V = X^2Y$.

7.28 Assume a product solution of Laplace's equation in cylindrical coordinates, $V = PF$, where V is not a function of z, P is a function only of ρ, and F is a function only of ϕ. (*a*) Obtain the two separated equations if the separation constant is n^2. Select the sign of n^2 so that the solution of the ϕ equation leads to trigonometric functions. (*b*) Show that $P = A\rho^n + B\rho^{-n}$ satisfies the ρ equation. (*c*) Construct the solution $V(\rho, \phi)$. Functions of this form are called *circular harmonics*.

CHAPTER

8

THE STEADY MAGNETIC FIELD

At this point the concept of a field should be a familiar one. Since we first accepted the experimental law of forces existing between two point charges and defined electric field intensity as the force per unit charge on a test charge in the presence of a second charge, we have discussed numerous fields. These fields possess no real physical basis, for physical measurements must always be in terms of the forces on the charges in the detection equipment. Those charges which are the source cause measurable forces to be exerted on other charges, which we may think of as detector charges. The fact that we attribute a field to the source charges and then determine the effect of this field on the detector charges amounts merely to a division of the basic problem into two parts for convenience.

We shall begin our study of the magnetic field with a definition of the magnetic field itself and show how it arsies from a current distribution. The effect of this field on other currents, or the second half of the physical problem, will be discussed in the following chapter. As we did with the electric field, we shall confine our initial discussion to free-space conditions, and the effect of material media will also be saved for discussion in the following chapter.

The relation of the steady magnetic field to its source is more complicated than is the relation of the electrostatic field to its source. We shall find it necessary to accept several laws temporarily on faith alone, relegating their proof to

the (rather difficult) final section in this chapter. This section may well be omitted when studying magnetic fields for the first time. It is included to make acceptance of the laws a little easier; the proof of the laws does exist and is available for the disbelievers or the more advanced student.

8.1 BIOT-SAVART LAW

The source of the steady magnetic field may be a permanent magnet, an electric field changing linearly with time, or a direct current. We shall largely ignore the permanent magnet and save the time-varying electric field for a later discussion. Our present relationships will concern the magnetic field produced by a differential dc element in free space.

We may think of this differential current element as a vanishingly small section of a current-carrying filamentary conductor, where a filamentary conductor is the limiting case of a cylindrical conductor of circular cross section as the radius approaches zero. We assume a current I flowing in a different vector length of the filament $d\mathbf{L}$. The law of Biot-Savart[1] then states that at any point P the magnitude of the magnetic field intensity produced by the differential element is proportional to the product of the current, the magnitude of the differential length, and the sine of the angle lying between the filament and a line connecting the filament to the point P at which the field is desired; also, the magnitude of the magnetic field intensity is inversely proportional to the square of the distance from the differential element to the point P. The direction of the magnetic field intensity is normal to the plane containing the differential filament and the line drawn from the filament to the point P. Of the two possible normals, that one is to be chosen which is in the direction of progress of a right-handed screw turned from $d\mathbf{L}$ through the smaller angle to the line from the filament to P. Using rationalized mks units, the constant of proportionality is $1/4\pi$.

The *Biot-Savart law*, described above in some 150 words, may be written concisely using vector notation as

$$d\mathbf{H} = \frac{I d\mathbf{L} \times \mathbf{a}_R}{4\pi R^2} = \frac{I d\mathbf{L} \times \mathbf{R}}{4\pi R^3} \tag{1}$$

The units of the *magnetic field intensity* $\mathbf{H}$ are evidently amperes per meter (A/m). The geometry is illustrated in Fig. 8.1. Subscripts may be used to indicate the point to which each of the quantities in (1) refers. If we locate the current element at point 1 and describe the point P at which the field is to be determined as point 2, then

[1] Biot and Savart were colleagues of Ampère, and all three were professors of physics at the Collège de France at one time or another. The Biot-Savart law was proposed in 1820.

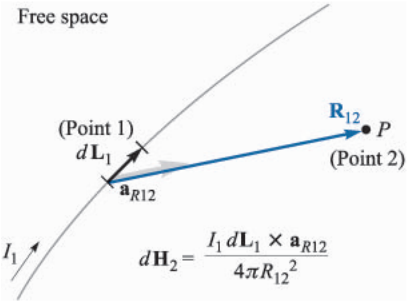


FIGURE 8.1
The law of Biot-Savart expresses the magnetic field intensity $d\mathbf{H}_2$ produced by a differential current element $I_1 d\mathbf{L}_1$. The direction of $d\mathbf{H}_2$ is into the page.

$$\boxed{d\mathbf{H}_2 = \frac{I_1 d\mathbf{L}_1 \times \mathbf{a}_{R12}}{4\pi R_{12}^2}} \tag{2}$$

The law of Biot-Savart is sometimes called *Ampère's law for the current element*, but we shall retain the former name because of possible confusion with Ampère's circuital law, to be discussed later.

In some aspects, the Biot-Savart law is reminiscent of Coulomb's law when that law is written for a differential element of charge,

$$d\mathbf{E}_2 = \frac{dQ_1 \mathbf{a}_{R12}}{4\pi\epsilon_0 R_{12}^2}$$

Both show an inverse-square-law dependence on distance, and both show a linear relationship between source and field. The chief difference appears in the direction of the field.

It is impossible to check experimentally the law of Biot-Savart as expressed by (1) or (2) because the differential current element cannot be isolated. We have restricted our attention to direct currents only, so the charge density is not a function of time. The continuity equation in Sec. 5.2, Eq. (5),

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho_v}{\partial t}$$

therefore shows that

$$\nabla \cdot \mathbf{J} = 0$$

or upon applying the divergence theorem,

$$\oint_s \mathbf{J} \cdot d\mathbf{S} = 0$$

The total current crossing any closed surface is zero, and this condition may be satisfied only by assuming a current flow around a closed path. It is this current flowing in a closed circuit which must be our experimental source, not the differential element.

It follows that only the integral form of the Biot-Savart law can be verified experimentally,

$$\mathbf{H} = \oint \frac{I d\mathbf{L} \times \mathbf{a}_R}{4\pi R^2} \tag{3}$$

Equation (1) or (2), of course leads directly to the integral form (3), but other differential expressions also yield the same integral formulation. Any term may be added to (1) whose integral around a closed path is zero. That is, any conservative field could be added to (1). The gradient of any scalar field always yields a conservative field, and we could therefore add a term ∇G to (1), where G is a general scalar field, without changing (3) in the slightest. This qualification on (1) or (2) is mentioned to show that if we later ask some foolish questions, not subject to any experimental check, concerning the force exerted by one *differential* current element on another, we should expect foolish answers.

The Biot-Savart law may also be expressed in terms of distributed sources, such as current density **J** and *surface current density* **K**. Surface current flows in a sheet of vanishingly small thickness, and the current density **J**, measured in amperes per square meter, is therefore infinite. Surface current density, however, is measured in amperes per meter width and designated by **K**. If the surface current density is uniform, the total current I in any width b is

$$I = Kb$$

where we have assumed that the width b is measured perpendicularly to the direction in which the current is flowing. The geometry is illustrated by Fig. 8.2. For a nonuniform surface current density, integration is necessary:

$$I = \int K dN \tag{4}$$

where dN is a differential element of the path *across* which the current is flowing. Thus the differential current element $I\,d\mathbf{L}$, where $d\mathbf{L}$ is in the direction of the

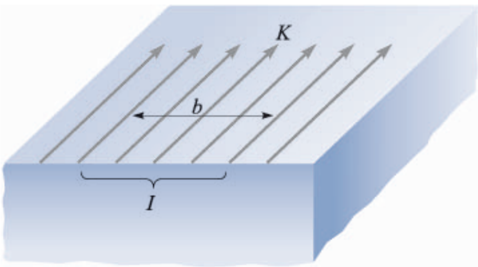


FIGURE 8.2
The total current I within a transverse width b, in which there is a *uniform* surface current density K, is Kb.

current, may be expressed in terms of surface current density **K** or current density **J**,

$$I\,d\mathbf{L} = \mathbf{K}\,dS = \mathbf{J}\,dv \qquad (5)$$

and alternate forms of the Biot-Savart law obtained,

$$\mathbf{H} = \int_s \frac{\mathbf{K} \times \mathbf{a}_R dS}{4\pi R^2} \qquad (6)$$

and

$$\mathbf{H} = \int_{\text{vol}} \frac{\mathbf{J} \times \mathbf{a}_R dv}{4\pi R^2} \qquad (7)$$

We may illustrate the application of the Biot-Savart law by considering an infinitely long straight filament. We shall apply (2) first and then integrate. This, of course, is the same as using the integral form (3) in the first place.[2]

Referring to Fig. 8.3, we should recognize the symmetry of this field. No variation with z or with ϕ can exist. Point 2, at which we shall determine the field, is therefore chosen in the $z = 0$ plane. The field point **r** is therefore $r = \rho\mathbf{a}_\rho$. The source point $\mathbf{r}'$ is given by $\mathbf{r}' = z'\mathbf{a}_z$, and therefore

$$\mathbf{R}_{12} = \mathbf{r} - \mathbf{r}' = \rho\mathbf{a}_\rho - z'\mathbf{a}_z$$

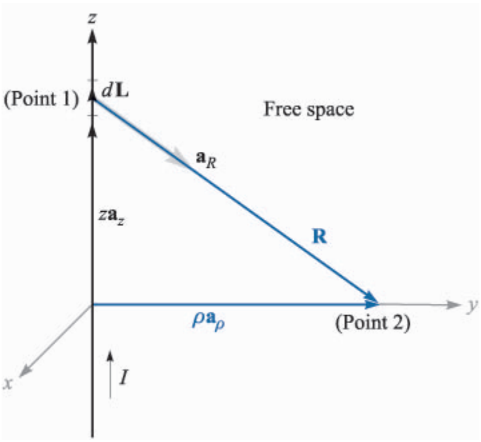


FIGURE 8.3
An infinitely long straight filament carrying a direct current I. The field at point 2 is $\mathbf{H} = (I/2\pi\rho)\mathbf{a}_\phi$.

[2] The closed path for the current may be considered to include a return filament parallel to the first filament and infinitely far removed. An outer coaxial conductor of infinite radius is another theoretical possibility. Practically, the problem is an impossible one, but we should realize that our answer will be quite accurate near a very long straight wire having a distant return path for the current.

so that

$$\mathbf{a}_{R12} = \frac{\rho \mathbf{a}_\rho - z'\mathbf{a}_z}{\sqrt{\rho^2 + z'^2}}$$

We take $d\mathbf{L} = dz'\mathbf{a}_z$ and (2) becomes

$$d\mathbf{H}_2 = \frac{I\, dz'\mathbf{a}_z \times (\rho\mathbf{a}_\rho - z'\mathbf{a}_z)}{4\pi(\rho^2 + z'^2)^{3/2}}$$

Since the current is directed toward increasing values of z', the limits are $-\infty$ and ∞ on the integral, and we have

$$\mathbf{H}_2 = \int_{-\infty}^{\infty} \frac{I\, dz'\mathbf{a}_z \times (\rho\mathbf{a}_\rho - z'\mathbf{a}_z)}{4\pi(\rho^2 + z'^2)^{3/2}}$$
$$= \frac{I}{4\pi}\int_{-\infty}^{\infty} \frac{\rho dz'\mathbf{a}_\phi}{(\rho^2 + z'^2)^{3/2}}$$

At this point the unit vector $\mathbf{a}_\phi$, under the integral sign should be investigated, for it is not always a constant, as are the unit vectors of the cartesian coordinate system. A vector is constant when its magnitude and direction are both constant. The unit vector certainly has constant magnitude, but its direction may change. Here $\mathbf{a}_\phi$ changes with the coordinate ϕ but not with ρ or z. Fortunately, the integration here is with respect to z', and $\mathbf{a}_\phi$ is a constant and may be removed from under the integral sign,

$$\mathbf{H}_2 = \frac{I\rho\mathbf{a}_\phi}{4\pi}\int_{-\infty}^{\infty} \frac{dz'}{(\rho^2 + z'^2)^{3/2}}$$
$$= \frac{I\rho\mathbf{a}_\phi}{4\pi} \left. \frac{z'}{\rho^2\sqrt{\rho^2 + z'^2}} \right|_{-\infty}^{\infty}$$

and

$$\boxed{\mathbf{H}_2 = \frac{I}{2\pi\rho}\mathbf{a}_\phi} \qquad (8)$$

The magnitude of the field is not a function of ϕ or z and it varies inversely as the distance from the filament. The direction of the magnetic-field-intensity vector is circumferential. The streamlines are therefore circles about the filament, and the field may be mapped in cross section as in Fig. 8.4.

The separation of the streamlines is proportional to the radius, or inversely proportional to the magnitude of **H**. To be specific, the streamlines have been drawn with curvilinear squares in mind. As yet we have no name for the family of lines[3] which are perpendicular to these circular streamlines, but the spacing of

[3] If you can't wait, see Sec. 8.6

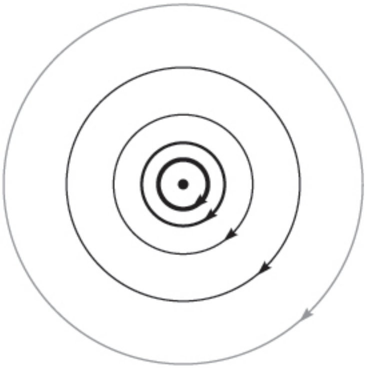


FIGURE 8.4
The streamlines of the magnetic field intensity about an infinitely long straight filament carrying a direct current I. The direction of I is into the page.

the streamlines has been adjusted so that the addition of this second set of lines will produce an array of curvilinear squares.

A comparison of Fig. 8.4 with the map of the *electric* field about an infinite line *charge* shows that the streamlines of the magnetic field correspond exactly to the equipotentials of the electric field, and the unnamed (and undrawn) perpendicular family of lines in the magnetic field corresponds to the streamlines of the electric field. This correspondence is not an accident, but there are several other concepts which must be mastered before the analogy between electric and magnetic fields can be explored more thoroughly.

Using the Biot-Savart law to find **H** is in many respects similar to the use of Coulomb's law to find **E**. Each requires the determination of a moderately complicated integrand containing vector quantities, followed by an integration. When we were concerned with Coulomb's law we solved a number of examples, including the fields of the point charge, line charge, and sheet of charge. The law of Biot-Savart can be used to solve analogous problems in magnetic fields, and some of these problems now appear as exercises at the end of the chapter rather than as examples here.

One useful result is the field of the finite-length current element, shown in Fig. 8.5. It turns out (see Prob. 8 at the end of the chapter) that **H** is most easily expressed in terms of the angles α_1 and α_2, as identified in the figure. The result is

$$\mathbf{H} = \frac{I}{4\pi\rho}(\sin\alpha_2 - \sin\alpha_1)\mathbf{a}_\phi \tag{9}$$

If one or both ends are below point 2, then α_1, or both α_1 and α_2, are negative.

Equation (9) may be used to find the magnetic field intensity caused by current filaments arranged as a sequence of straight line segments.

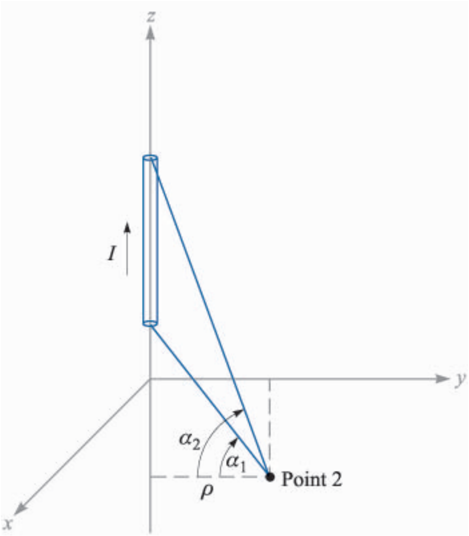


FIGURE 8.5
The magnetic field intensity caused by a finite-length current filament on the z axis is $(I/4\pi\rho)(\sin\alpha_2 - \sin\alpha_1)\mathbf{a}_\phi$.

Example 8.1

As a numerical example illustrating the use of (9), let us determine **H** at $P_2(0.4, 0.3, 0)$ in the field of an 8-A filamentary current directed inward from infinity to the origin on the positive x axis, and then outward to infinity along the y axis. This arrangement is shown in Figure 8.6.

Solution. We first consider the semi-infinite current on the x axis, identifying the two angles, $\alpha_{1x} = -90^\circ$ and $\alpha_{2x} = \tan^{-1}(0.4/0.3) = 53.1^\circ$. The radial distance ρ is measured from the x axis, and we have $\rho_x = 0.3$. Thus, this contribution to $\mathbf{H}_2$ is

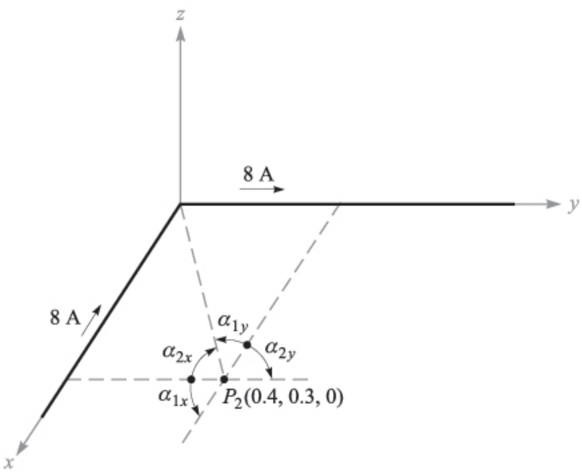


FIGURE 8.6
The individual fields of two semi-infinite current segments are found by (9) and added to obtain $\mathbf{H}_2$ at P_2.

$$\mathbf{H}_{2(x)} = \frac{8}{4\pi(0.3)}(\sin 53.1^\circ + 1)\mathbf{a}_\phi = \frac{2}{0.3\pi}(1.8)\mathbf{a}_\phi = \frac{12}{\pi}\mathbf{a}_\phi$$

The unit vector $\mathbf{a}_\phi$, must also be referred to the x axis. We see that it becomes $-\mathbf{a}_z$. Therefore,

$$\mathbf{H}_{2(x)} = -\frac{12}{\pi}\mathbf{a}_z \quad \text{A/m}$$

For the current on the y axis, we have $\alpha_{1y} = -\tan^{-1}(0.3/0.4) = -36.9^\circ$, $\alpha_{2y} = 90^\circ$, and $\rho_y = 0.4$. It follows that

$$\mathbf{H}_{2(y)} = \frac{8}{4\pi(0.4)}(1 + \sin 36.9^\circ)(-\mathbf{a}_z) = -\frac{8}{\pi}\mathbf{a}_z \quad \text{A/m}$$

Adding these results, we have

$$\mathbf{H}_2 = \mathbf{H}_{2(x)} + \mathbf{H}_{2(y)} = -\frac{20}{\pi}\mathbf{a}_z = -6.37\mathbf{a}_z \quad \text{A/m}$$

✔ **D8.1.** Given the following values for P_1, P_2, and $I_1\Delta_1$, calculate $\Delta\mathbf{H}_2$: (*a*) $P_1(0,0,2)$, $P_2(4,2,0)$, $2\pi\mathbf{a}_z \mu\text{A}\cdot\text{m}$; (*b*) $P_1(0,2,0)$, $P_2(4,2,0)$, $2\pi\mathbf{a}_z \mu\text{A}\cdot\text{m}$; (*c*) $P_1(1,2,3)$, $P_2(-3,-1,2)$, $2\pi(-\mathbf{a}_x + \mathbf{a}_y + 2\mathbf{a}_z)\mu\text{A}\cdot\text{m}$.

Ans. $-8.51\mathbf{a}_x + 17.01\mathbf{a}_y$ nA/m; $16\mathbf{a}_y$ nA/m; $3.77\mathbf{a}_x - 6.79\mathbf{a}_y + 5.28\mathbf{a}_z$ nA/m

✔ **D8.2.** A current filament carrying 15 A in the $\mathbf{a}_z$ direction lies along the entire z axis. Find **H** in cartesian coordinates at: (*a*) $P_A(\sqrt{20}, 0, 4)$; (*b*) $P_B(2, -4, 4)$.

Ans. $0.534\mathbf{a}_y$ A/m; $0.477\mathbf{a}_x + 0.239\mathbf{a}_y$ A/m

8.2 AMPÈRE'S CIRCUITAL LAW

After solving a number of simple electrostatic problems with Coulomb's law, we found that the same problems could be solved much more easily by using Gauss's law whenever a high degree of symmetry was present. Again, an analogous procedure exists in magnetic fields. Here, the law that helps us solve problems more easily is known as *Ampère's circuital*[4] *law*, sometimes called Ampère's work law. This law may be derived from the Biot-Savart law, and the derivation is accomplished in Sec. 8.7. For the present we might agree to accept Ampère's circuital law temporarily as another law capable of experimental proof. As is the case with Gauss's law, its use will also require careful consideration of the symmetry of the problem to determine which variables and components are present.

Ampère's circuital law states that the line integral of **H** about any *closed* path is exactly equal to the direct current enclosed by that path,

[4] The preferred pronunciation puts the accent on "circ-."

$$\oint \mathbf{H} \cdot d\mathbf{L} = I \qquad (10)$$

We define positive current as flowing in the direction of advance of a right-handed screw turned in the direction in which the closed path is traversed.

Referring to Fig. 8.7, which shows a circular wire carrying a direct current I, the line integral of $\mathbf{H}$ about the closed paths lettered a and b results in an answer of I; the integral about the closed path c which passes through the conductor gives an answer less than I and is exactly that portion of the total current which is enclosed by the path c. Although paths a and b give the same answer, the integrands are, of course, different. The line integral directs us to multiply the component of $\mathbf{H}$ in the direction of the path by a small increment of path length at one point of the path, move along the path to the next incremental length, and repeat the process, continuing until the path is completely traversed. Since $\mathbf{H}$ will generally vary from point to point, and since paths a and b are not alike, the contributions to the integral made by, say, each micrometer of path length are quite different. Only the final answers are the same.

We should also consider exactly what is meant by the expression "current enclosed by the path." Suppose we solder a circuit together after passing the conductor once through a rubber band, which we shall use to represent the closed path. Some strange and formidable paths can be constructed by twisting and knotting the rubber band, but if neither the rubber band nor the conducting circuit is broken, the current enclosed by the path is that carried by the conductor. Now let us replace the rubber band by a circular ring of spring steel across which is stretched a rubber sheet. The steel loop forms the closed path, and the current-carrying conductor must pierce the rubber sheet if the current is to be enclosed by the path. Again, we may twist the steel loop, and we may also deform the rubber sheet by pushing our fist into it or folding it in any way we wish. A single current-carrying conductor still pierces the sheet once, and this is the true measure of the current enclosed by the path. If we should thread the conductor once through the sheet from front to back and once from back to front, the total current enclosed by the path is the algebraic sum, which is zero.

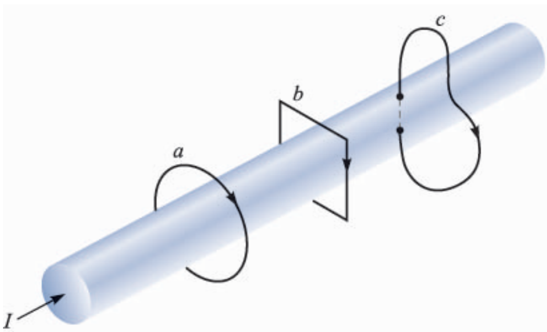


FIGURE 8.7
A conductor has a total current I. The line integral of $\mathbf{H}$ about the closed paths a and b is equal to I, and the integral around path c is less than I, since the entire current is not enclosed by the path.

In more general language, given a closed path, we recognize this path as the perimeter of an infinite number of surfaces (not closed surfaces). Any current-carrying conductor enclosed by the path must pass through every one of these surfaces once. Certainly some of the surfaces may be chosen in such a way that the conductor pierces them twice in one direction and once in the other direction, but the algebraic total current is still the same.

We shall find that the nature of the closed path is usually extremely simple and can be drawn on a plane. The simplest surface is, then, that portion of the plane enclosed by the path. We need merely find the total current passing through this region of the plane.

The application of Gauss's law involves finding the total charge enclosed by a closed surface; the application of Ampère's circuital law involves finding the total current enclosed by a closed path.

Let us again find the magnetic field intensity produced by an infinitely long filament carrying a current I. The filament lies on the z axis in free space (as in Fig. 8.3), and the current flows in the direction given by $\mathbf{a}_z$. Symmetry inspection comes first, showing that there is no variation with z or ϕ. Next we determine which components of $\mathbf{H}$ are present by using the Biot-Savart law. Without specifically using the cross product, we may say that the direction of $d\mathbf{H}$ is perpendicular to the plane conaining $d\mathbf{L}$ and $\mathbf{R}$ and therefore is in the direction of $\mathbf{a}_\phi$. Hence the only component of $\mathbf{H}$ is H_ϕ, and it is a function only of ρ.

We therefore choose a path to any section of which $\mathbf{H}$ is either perpendicular or tangential and along which H is constant. The first requirement (perpendicularity or tangency) allows us to replace the dot product of Ampère's circuital law with the product of the scalar magnitudes, except along that portion of the path where $\mathbf{H}$ is normal to the path and the dot product is zero; the second requirement (constancy) then permits us to remove the magnetic field intensity from the integral sign. The integration required is usually trivial and consists of finding the length of that portion of the path to which $\mathbf{H}$ is parallel.

In our example the path must be a circle of radius ρ and Ampère's circuital law becomes

$$\oint \mathbf{H} \cdot d\mathbf{L} = \int_0^{2\pi} H_\phi \rho d\phi = H_\phi \rho \int_0^{2\pi} d\phi = H_\phi 2\pi\rho = I$$

or

$$H_\phi = \frac{I}{2\pi\rho}$$

as before.

As a second example of the application of Ampère's circuital law, consider an infinitely long coaxial transmission line carrying a uniformly distributed total current I in the center conductor and $-I$ in the outer conductor. The line is shown in Fig. 8.8*a*. Symmetry shows that H is not a function of ϕ or z. In order to determine the components present, we may use the results of the previous

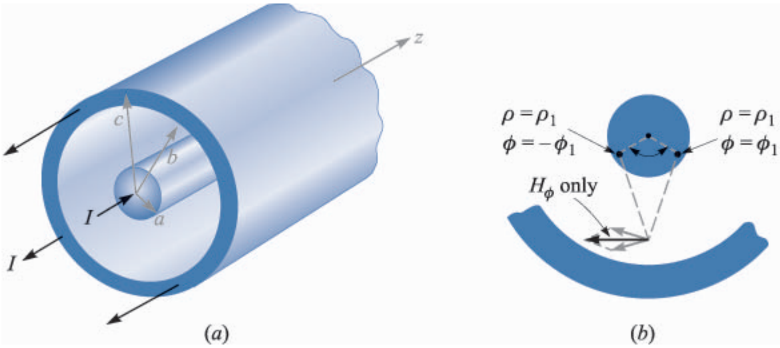


FIGURE 8.8
(*a*) Cross section of a coaxial cable carrying a uniformly distributed current I in the inner conductor and $-I$ in the outer conductor. The magnetic field at any point is most easily determined by applying Ampère's circuital law about a circular path. (*b*) Current filaments at $\rho = \rho_1$, $\phi = \pm\phi_1$, produces $\mathbf{H}_\rho$ components which cancel. For the total field, $\mathbf{H} = \mathbf{H}_\phi \mathbf{a}_\phi$.

example by considering the solid conductors as being composed of a large number of filaments. No filament has a z component of $\mathbf{H}$. Furthermore, the H_ρ component at $\phi = 0°$, produced by one filament located at $\rho = \rho_1$, $\phi = \phi_1$, is canceled by the H_ρ component produced by a symmetrically located filament at $\rho = \rho_1$, $\phi = -\phi_1$. This symmetry is illustrated by Fig. 8.8*b*. Again we find only an H_ϕ component which varies with ρ.

A circular path of radius ρ, where ρ is larger than the radius of the inner conductor but less than the inner radius of the outer conductor, then leads immediately to

$$H_\phi = \frac{I}{2\pi\rho} \quad (a < \rho < b)$$

If we choose ρ smaller than the radius of the inner conductor, the current enclosed is

$$I_{encl} = I\frac{\rho^2}{a^2}$$

and

$$2\pi\rho H_\phi = I\frac{\rho^2}{a^2}$$

or

$$H_\phi = \frac{I\rho}{2\pi a^2} \quad (\rho < a)$$

If the radius ρ is larger than the outer radius of the outer conductor, no current is enclosed and

$$H_\phi = 0 \quad (\rho > c)$$

Finally, if the path lies within the outer conductor, we have

$$2\pi\rho H_\phi = I - I\left(\frac{\rho^2 - b^2}{c^2 - b^2}\right)$$

$$H_\phi = \frac{I}{2\pi\rho}\frac{c^2 - \rho^2}{c^2 - b^2} \quad (b < \rho < c)$$

The magnetic-field-strength variation with radius is shown in Fig. 8.9 for a coaxial cable in which $b = 3a$, $c = 4a$. It should be noted that the magnetic field intensity **H** is continuous at all the conductor boundaries. In other words, a slight increase in the radius of the closed path does not result in the enclosure of a tremendously different current. The value of H_ϕ shows no sudden jumps.

The external field is zero. This, we see, results from equal positive and negative currents enclosed by the path. Each produces an external field of magnitude $I/2\pi\rho$, but complete cancellation occurs. This is another example of "shielding"; such a coaxial cable carrying large currents would not produce any noticeable effect in an adjacent circuit.

As a final example, let us consider a sheet of current flowing in the positive y direction and located in the $z = 0$ plane. We may think of the return current as equally divided between two distant sheets on either side of the sheet we are considering. A sheet of uniform surface current density $\mathbf{K} = K_y\mathbf{a}_y$ is shown in Fig. 8.10. **H** cannot vary with x or y. If the sheet is subdivided into a number of filaments, it is evident that no filament can produce an H_y component. Moreover, the Biot-Savart law shows that the contributions to H_z produced by a symmetrically located pair of filaments cancel. Thus, H_z is zero also; only an H_x component is present. We therefore choose the path 1-1′-2′-2-1 composed of straight-line segments which are either parallel or perpendicular to H_x. Ampère's circuital law gives

$$H_{x1}L + H_{x2}(-L) = K_yL$$

or

$$H_{x1} - H_{x2} = K_y$$

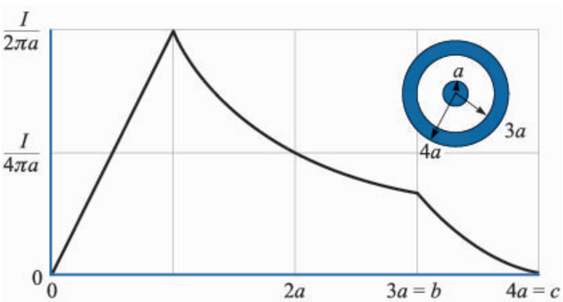


FIGURE 8.9
The magnetic field intensity as a function of radius in an infinitely long coaxial transmission line with the dimensions shown.

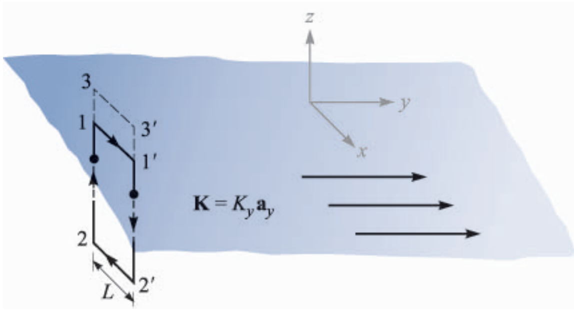


FIGURE 8.10
A uniform sheet of surface current $\mathbf{K} = K_y \mathbf{a}_y$ in the $z = 0$ plane. $\mathbf{H}$ may be found by applying Ampère's circuital law about the paths 1-1′-2′-2-1 and 3-3′-2′-2-3.

If the path 3-3′-2′-2-3 is now chosen, the same current is enclosed, and

$$H_{x3} - H_{x2} = K_y$$

and therefore

$$H_{x3} = H_{x1}$$

It follows that H_x is the same for all positive z. Similarly, H_x is the same for all negative z. Because of the symmetry, then, the magnetic field intensity on one side of the current sheet is the negative of that on the other. Above the sheet,

$$H_x = \tfrac{1}{2} K_y \quad (z > 0)$$

while below it

$$H_x = -\tfrac{1}{2} K_y \quad (z < 0)$$

Letting $\mathbf{a}_N$ be a unit vector normal (outward) to the current sheet, the result may be written in a form correct for all z as

$$\boxed{\mathbf{H} = \tfrac{1}{2} \mathbf{K} \times \mathbf{a}_N} \tag{11}$$

If a second sheet of current flowing in the opposite direction, $\mathbf{K} = -K_y \mathbf{a}_y$, is placed at $z = h$, (11) shows that the field in the region between the current sheets is

$$\boxed{\mathbf{H} = \mathbf{K} \times \mathbf{a}_N \quad (0 < z < h)} \tag{12}$$

and is zero elsewhere,

$$\boxed{\mathbf{H} = 0 \quad (z < 0, z > h)} \tag{13}$$

The most difficult part of the application of Ampère's circuital law is the determination of the components of the field which are present. The surest

method is the logical application of the Biot-Savart law and a knowledge of the magnetic fields of simple form.

Problem 13 at the end of this chapter outlines the steps involved in applying Ampère's circuital law to an infinitely long solenoid of radius a and uniform current density $K_a\mathbf{a}_\phi$, as shown in Fig. 8.11*a*. For reference, the result is

$$\mathbf{H} = K_a\mathbf{a}_z \quad (\rho < a) \tag{14a}$$

$$\mathbf{H} = 0 \quad (\rho > a) \tag{14b}$$

If the solenoid has a finite length d and consists of N closely wound turns of a filament that carries a current I (Fig. 8.11b), then the field at points well within the solenoid is given closely by

$$\mathbf{H} = \frac{NI}{d}\mathbf{a}_z \quad \text{(well within the solenoid)} \tag{15}$$

The approximation is useful it if is not applied closer than two radii to the open ends, nor closer to the solenoid surface than twice the separation between turns.

For the toroids shown in Fig. 8.12, it can be shown that the magnetic field intensity for the ideal case, Fig. 8.12*a*, is

$$\mathbf{H} = K_a\frac{\rho_0 - a}{\rho}\mathbf{a}_\phi \quad \text{(inside toroid)} \tag{16a}$$

$$\mathbf{H} = 0 \quad \text{(outside)} \tag{16b}$$

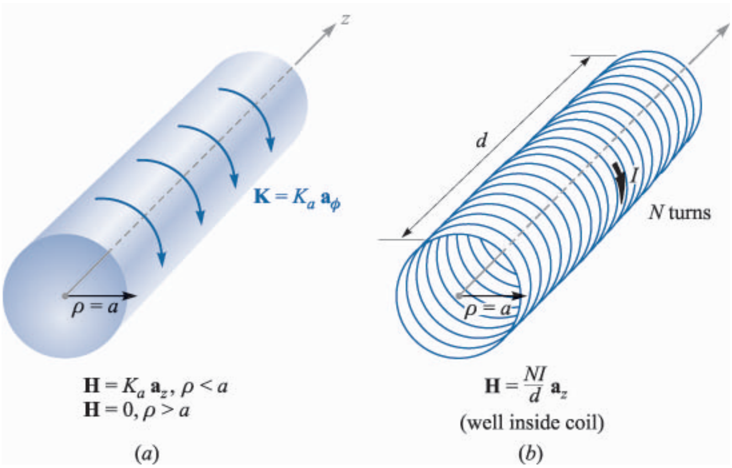


FIGURE 8.11
(*a*) An ideal solenoid of infinite length with a circular current sheet $\mathbf{K} = K_a\mathbf{a}_\phi$. (*b*) An N-turn solenoid of finite length d.

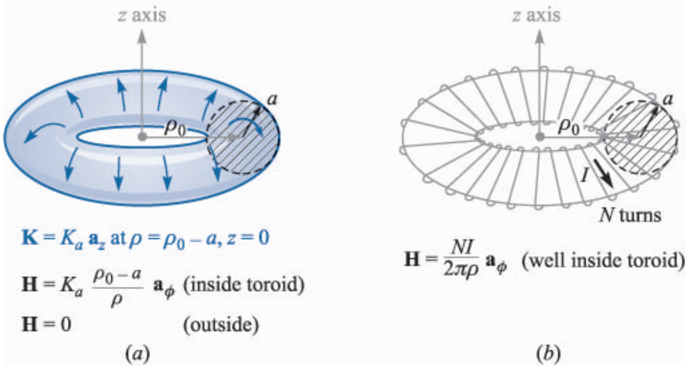


FIGURE 8.12
(*a*) An ideal toroid carrying a surface current **K** in the direction shown. (*b*) An N-turn toroid carrying a filamentary current I.

For the N-turn toroid of Figure 8.12*b*, we have the good approximations,

$$\mathbf{H} = \frac{NI}{2\pi\rho}\mathbf{a}_\phi \quad \text{(inside toroid)} \tag{17a}$$

$$\mathbf{H} = 0 \quad \text{(outside)} \tag{17b}$$

as long as we consider points removed from the toroidal surface by several times the separation between turns.

Toroids having rectangular cross sections are also treated quite readily, as you can see for yourself by trying Prob. 14.

Accurate formulas for solenoids, toroids, and coils of other shapes are available in Sec. 2 of the "Standard Handbook for Electrical Engineers" (see Suggested References for Chap. 5).

✔ **D8.3.** Express the value of **H** in cartesian components at $P(0, 0.2, 0)$ in the field of: (*a*) a current filament, 2.5 A in the $\mathbf{a}_z$ direction at $x = 0.1$, $y = 0.3$; (*b*) a coax, centered on the z axis, with $a = 0.3$, $b = 0.5$, $c = 0.6$, $I = 2.5$ A in $\mathbf{a}_z$ direction in center conductor; (*c*) three current sheets, $2.7\mathbf{a}_x$ A/m at $y = 0.1$, $-1.4\mathbf{a}_x$ A/m at $y = 0.15$, and $-1.3\mathbf{a}_x$ A/m at $y = 0.25$.

Ans. $1.989\mathbf{a}_x - 1.989\mathbf{a}_y$ A/m; $-0.884\mathbf{a}_x$ A/m; $1.300\mathbf{a}_z$ A/m

8.3 CURL

We completed our study of Gauss's law by applying it to a differential volume element and were led to the concept of divergence. We now apply Ampère's circuital law to the perimeter of a differential surface element and discuss the

third and last of the special derivatives of vector analysis, the curl. Our immediate objective is to obtain the point form of Ampère's circuital law.

Again we shall choose cartesian coordinates, and an incremental closed path of sides Δx and Δy is selected (Fig. 8.13). We assume that some current, as yet unspecified, produces a reference value for **H** at the *center* of this small rectangle,

$$\mathbf{H}_0 = H_{x0}\mathbf{a}_x + H_{y0}\mathbf{a}_y + H_{z0}\mathbf{a}_z$$

The closed line integral of **H** about this path is then approximately the sum of the four values of $\mathbf{H} \cdot \Delta\mathbf{L}$ on each side. We choose the direction of traverse as 1-2-3-4-1, which corresponds to a current in the $\mathbf{a}_z$ direction, and the first contribution is therefore

$$(\mathbf{H} \cdot \Delta\mathbf{L})_{1-2} = H_{y,1-2}\Delta y$$

The value of H_y *on this section* of the path may be given in terms of the reference value H_{y0} at the center of the rectangle, the rate of change of H_y with x, and the distance $\Delta x/2$ from the center to the midpoint of side 1-2:

$$H_{y,1-2} \doteq H_{y0} + \frac{\partial H_y}{\partial x}\left(\tfrac{1}{2}\Delta x\right)$$

Thus

$$(\mathbf{H} \cdot \Delta\mathbf{L})_{1-2} \doteq \left(H_{y0} + \frac{1}{2}\frac{\partial H_y}{\partial x}\Delta x\right)\Delta y$$

Along the next section of the path we have

$$(\mathbf{H} \cdot \Delta\mathbf{L})_{2-3} \doteq H_{x,2-3}(-\Delta x) \doteq -\left(H_{x0} + \frac{1}{2}\frac{\partial H_x}{\partial y}\Delta y\right)\Delta x$$

Continuing for the remaining two segments and adding the results,

$$\oint \mathbf{H} \cdot d\mathbf{L} \doteq \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\Delta x \Delta y$$

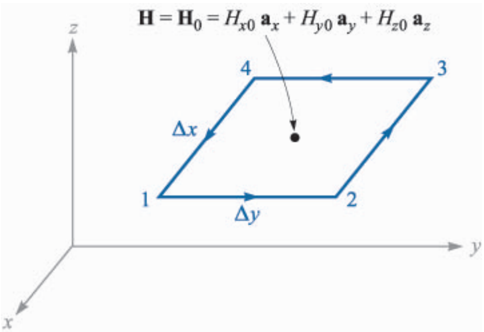


FIGURE 8.13
An incremental closed path in cartesian coordinates is selected for the application of Ampère's circuital law to determine the spatial rate of change of **H**.

By Ampère's circuital law, this result must be equal to the current enclosed by the path, or the current crossing any surface bounded by the path. If we assume a general current density **J**, the enclosed current is then $\Delta I \doteq J_z \Delta x \Delta y$, and

$$\oint \mathbf{H} \cdot d\mathbf{L} \doteq \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right) \Delta x \Delta y \doteq J_z \Delta x \Delta y$$

or

$$\frac{\oint \mathbf{H} \cdot d\mathbf{L}}{\Delta x \Delta y} \doteq \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \doteq J_z$$

As we cause the closed path to shrink, the above expression becomes more nearly exact, and in the limit we have the equality

$$\lim_{\Delta x, \Delta y \to 0} \frac{\oint \mathbf{H} \cdot d\mathbf{L}}{\Delta x \Delta y} = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = J_z \tag{18}$$

After beginning with Ampère's circuital law equating the closed line integral of **H** to the current enclosed, we have now arrived at a relationship involving the closed line integral of **H** *per unit area* enclosed and the current *per unit area* enclosed, or current density. We performed a similar analysis in passing from the integral form of Gauss's law, involving flux through a closed surface and charge enclosed, to the point form, relating flux through a closed surface *per unit volume* enclosed and charge *per unit volume* enclosed, or volume charge density. In each case a limit is necessary to produce an equality.

If we choose closed paths which are oriented perpendicularly to each of the remaining two coordinate axes, analogous processes lead to expressions for the y and z components of the current density,

$$\lim_{\Delta y, \Delta z \to 0} \frac{\oint \mathbf{H} \cdot d\mathbf{L}}{\Delta y \Delta z} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = J_x \tag{19}$$

and

$$\lim_{\Delta z, \Delta x \to 0} \frac{\oint \mathbf{H} \cdot d\mathbf{L}}{\Delta z \Delta x} = \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = J_y \tag{20}$$

Comparing (18), (19), and (20), we see that a component of the current density is given by the limit of the quotient of the closed line integral of **H** about a small path in a plane normal to that component and of the area enclosed as the path shrinks to zero. This limit has its counterpart in other fields of science and long ago received the name of *curl*. The curl of any vector is a vector, and any component of the curl is given by the limit of the quotient of the closed line integral of the vector about a small path in a plane normal to that component desired and the area enclosed, as the path shrinks to zero. It should be noted that

the above definition of curl does not refer specifically to a particular coordinate system. The mathematical form of the definition is

$$(\text{curl } \mathbf{H})_N = \lim_{\Delta S_N \to 0} \frac{\oint \mathbf{H} \cdot d\mathbf{L}}{\Delta S_N} \tag{21}$$

where ΔS_N is the planar area enclosed by the closed line integral. The N subscript indicates that the component of the curl is that component which is *normal* to the surface enclosed by the closed path. It may represent any component in any coordinate system.

In cartesian coordinates the definition (21) shows that the x, y, and z components of the curl **H** are given by (18), (19), and (20), and therefore

$$\text{curl } \mathbf{H} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right)\mathbf{a}_x + \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right)\mathbf{a}_y + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\mathbf{a}_z \tag{22}$$

This result may be written in the form of a determinant,

$$\text{curl } \mathbf{H} = \begin{vmatrix} \mathbf{a}_x & \mathbf{a}_y & \mathbf{a}_z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ H_x & H_y & H_z \end{vmatrix} \tag{23}$$

and may also be written in terms of the vector operator,

$$\text{curl } \mathbf{H} = \nabla \times \mathbf{H} \tag{24}$$

Equation (22) is the result of applying the definition (21) to the cartesian coordinate system. We obtained the z component of this expression by evaluating Ampère's circuital law about an incremental path of sides Δx and Δy, and we could have obtained the other two components just as easily by choosing the appropriate paths. Equation (23) is a neat method of storing the cartesian coordinate expression for curl; the form is symmetrical and easily remembered. Equation (24) is even more concise and leads to (22) upon applying the definitions of the cross product and vector operator.

The expressions for curl **H** in cylindrical and spherical coordinates are derived in Appendix A by applying the definition (21). Although they may be written in determinant form, as explained there, the determinants do not have one row of unit vectors on top and one row of components on the bottom, and they are not easily memorized. For this reason, the curl expansions in cylindrical

and spherical coordinates which appear below and inside the back cover are usually referred to whenever necessary.

$$\nabla \times \mathbf{H} = \left(\frac{1}{\rho}\frac{\partial H_z}{\partial \phi} - \frac{\partial H_\phi}{\partial z}\right)\mathbf{a}_\rho + \left(\frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho}\right)\mathbf{a}_\phi + \left(\frac{1}{\rho}\frac{\partial(\rho H_\phi)}{\partial \rho} - \frac{1}{\rho}\frac{\partial H_\rho}{\partial \phi}\right)\mathbf{a}_z \quad \text{(cylindrical)} \tag{25}$$

$$\nabla \times \mathbf{H} = \frac{1}{r\sin\theta}\left(\frac{\partial(H_\phi \sin\theta)}{\partial \theta} - \frac{\partial H_\theta}{\partial \phi}\right)\mathbf{a}_r + \frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial H_r}{\partial \phi} - \frac{\partial(rH_\phi)}{\partial r}\right)\mathbf{a}_\theta + \frac{1}{r}\left(\frac{\partial(rH_\theta)}{\partial r} - \frac{\partial H_r}{\partial \theta}\right)\mathbf{a}_\phi \quad \text{(spherical)} \tag{26}$$

Although we have described curl as a line integral per unit area, this does not provide everyone with a satisfactory physical picture of the nature of the curl operation, for the closed line integral itself requires physical interpretation. This integral was first met in the electrostatic field, where we saw that $\oint \mathbf{E} \cdot d\mathbf{L} = 0$. Inasmuch as the integral was zero, we did not belabor the physical picture. More recently we have discussed the closed line integral of **H**, $\oint \mathbf{H} \cdot d\mathbf{L} = I$. Either of these closed line integrals is also known by the name of "circulation," a term obviously borrowed from the field of fluid dynamics.

The circulation of **H**, or $\oint \mathbf{H} \cdot d\mathbf{L}$, is obtained by multiplying the component of **H** parallel to the specified closed path at each point along it by the differential path length and summing the results as the differential lengths approach zero and as their number becomes infinite. We do not require a vanishingly small path. Ampère's circuital law tells us that if **H** does possess circulation about a given path, then current passes through this path. In electrostatics we see that the circulation of **E** is zero about every path, a direct consequence of the fact that zero work is required to carry a charge around a closed path.

We may now describe curl as *circulation per unit area*. The closed path is vanishingly small, and curl is defined at a point. The curl of **E** must be zero, for the circulation is zero. The curl of **H** is not zero, however; the circulation of **H** per unit area is the current density by Ampère's circuital law [or (18), (19), and (20)].

Skilling[5] suggests the use of a very small paddle wheel as a "curl meter." Our vector quantity, then, must be thought of as capable of applying a force to each blade of the paddle wheel, the force being proportional to the component of the field normal to the surface of that blade. To test a field for curl we dip our paddle wheel into the field, with the axis of the paddle wheel lined up with the

[5] See the Suggested References at the end of the chapter.

direction of the component of curl desired, and note the action of the field on the paddle. No rotation means no curl; larger angular velocities mean greater values of the curl; a reversal in the direction of spin means a reversal in the sign of the curl. To find the direction of the vector curl and not merely to establish the presence of any particular component, we should place our paddle wheel in the field and hunt around for the orientation which produces the greatest torque. The direction of the curl is then along the axis of the paddle wheel, as given by the right-hand rule.

As an example, consider the flow of water in a river. Fig. 8.14*a* shows the longitudinal section of a wide river taken at the middle of the river. The water velocity is zero at the bottom and increases linearly as the surface is approached. A paddle wheel placed in the position shown, with its axis perpendicular to the paper, will turn in a clockwise direction, showing the presence of a component of curl in the direction of an inward normal to the surface of the page. If the velocity of water does not change as we go up- or downstream and also shows no variation as we go across the river (or even if it decreases in the same fashion toward either bank), then this component is the only component present at the center of the stream, and the curl of the water velocity has a direction into the page.

In Fig. 8.14*b* the streamlines of the magnetic field intensity about an infinitely long filamentary conductor are shown. The curl meter placed in this field of curved lines shows that a larger number of blades have a clockwise force exerted on them but that this force is in general smaller than the counterclockwise force exerted on the smaller number of blades closer to the wire. It seems possible that if the curvature of the streamlines is correct and also if the variation of the field strength is just right, the net torque on the paddle wheel may be zero. Actually, the paddle wheel does not rotate in this case, for since $\mathbf{H} = (I/2\pi\rho)\mathbf{a}_\phi$, we may substitute into (25) obtaining

$$\text{curl } \mathbf{H} = -\frac{\partial H_\phi}{\partial z}\mathbf{a}_\rho + \frac{1}{\rho}\frac{\partial(\rho H_\phi)}{\partial \rho}\mathbf{a}_z = 0$$

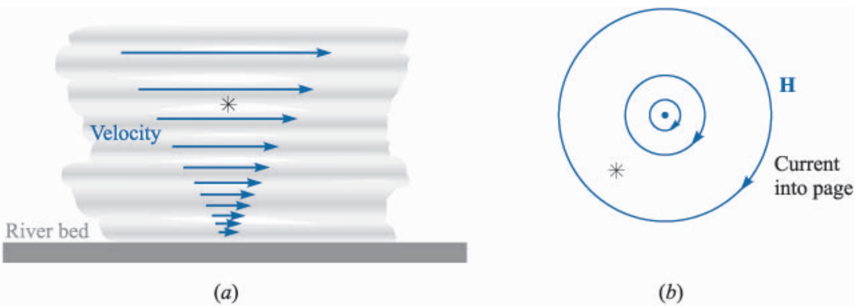


FIGURE 8.14
(*a*) The curl meter shows a component of the curl of the water velocity into the page. (*b*) The curl of the magnetic field intensity about an infinitely long filament is shown.

Example 8.2

As an example of the evaluation of curl **H** from the definition and of the evaluation of another line integral, let us suppose that $\mathbf{H} = 0.2z^2\mathbf{a}_x$ for $z > 0$, and $\mathbf{H} = 0$ elsewhere, as shown in Fig. 8.15. Calculate $\oint \mathbf{H} \cdot d\mathbf{L}$ about a square path with side d, centered at $(0, 0, z_1)$ in the $y = 0$ plane where $z_1 > 2d$.

Solution. We evaluate the line integral of **H** along the four segments, beginning at the top:

$$\oint \mathbf{H} \cdot d\mathbf{L} = 0.2(z_1 + \tfrac{1}{2}d)^2 d + 0 - 0.2(z_1 - \tfrac{1}{2}d)^2 d + 0$$
$$= 0.4 z_1 d^2$$

In the limit as the area approaches zero, we find

$$(\nabla \times \mathbf{H})_y = \lim_{d \to 0} \frac{\oint \mathbf{H} \cdot d\mathbf{L}}{d^2} = \lim_{d \to 0} \frac{0.4 z_1 d^2}{d^2} = 0.4 z_1$$

The other components are zero, so $\nabla \times \mathbf{H} = 0.4 z_1 \mathbf{a}_y$.

To evaluate the curl without trying to illustrate the definition or the evaluation of a line integral, we simply take the partial derivative indicated by (23):

$$\nabla \times \mathbf{H} = \begin{vmatrix} \mathbf{a}_x & \mathbf{a}_y & \mathbf{a}_z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 0.2z^2 & 0 & 0 \end{vmatrix} = \frac{\partial}{\partial z}(0.2z^2)\mathbf{a}_y = 0.4z\mathbf{a}_y$$

which checks with the result above when $z = z_1$.

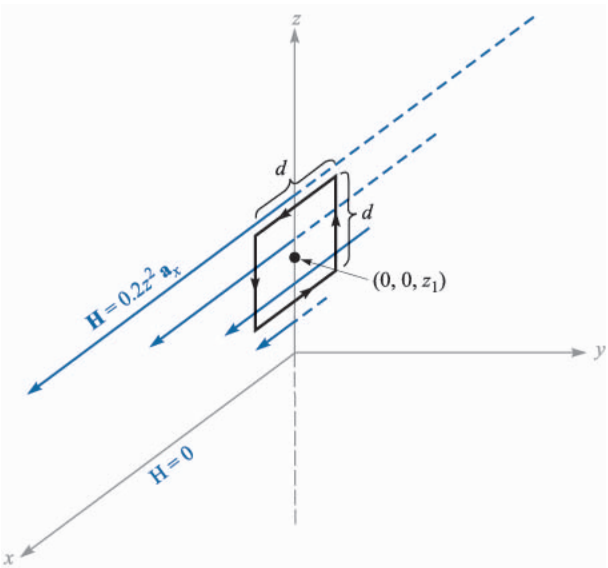


FIGURE 8.15
A square path of side d with its center on the z axis at $z = z_1$ is used to evaluate $\oint \mathbf{H} \cdot d\mathbf{L}$ and find curl **H**.

Returning now to complete our original examination of the application of Ampère's circuital law to a differential-sized path, we may combine (18), (19), (20), (22), and (24),

$$\text{curl } \mathbf{H} = \nabla \times \mathbf{H} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right)\mathbf{a}_x + \left(\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}\right)\mathbf{a}_y + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\mathbf{a}_z = \mathbf{J} \tag{27}$$

and write the *point form of Ampère's circuital law*,

$$\boxed{\nabla \times \mathbf{H} = \mathbf{J}} \tag{28}$$

This is the second of Maxwell's four equations as they apply to non-time-varying conditions. We may also write the third of these equations at this time; it is the point form of $\oint \mathbf{E} \cdot d\mathbf{L} = 0$, or

$$\boxed{\nabla \times \mathbf{E} = 0} \tag{29}$$

The fourth equation appears in Sec. 8.5.

✔ **D8.4.** (*a*) Evaluate the closed line integral of $\mathbf{H}$ about the rectangular path $P_1(2, 3, 4)$ to $P_2(4, 3, 4)$ to $P_3(4, 3, 1)$ to $P_4(2, 3, 1)$ to P_1, given $\mathbf{H} = 3z\mathbf{a}_x - 2x^3\mathbf{a}_z$ A/m. (*b*) Determine the quotient of the closed line integral and the area enclosed by the path as an approximation to $(\nabla \times \mathbf{H})_y$. (*c*) Determine $(\nabla \times \mathbf{H})_y$ at the center of the area.

Ans. 354 A; 59 A/m^2; 57 A/m^2

✔ **D8.5.** Calculate the value of the vector current density: (*a*) in cartesian coordinates at $P_A(2, 3, 4)$ if $\mathbf{H} = x^2 z\mathbf{a}_y - y^2 x\mathbf{a}_z$; (*b*) in cylindrical coordinates at $P_B(1.5, 90°, 0.5)$ if $\mathbf{H} = \frac{2}{\rho}(\cos 0.2\phi)\mathbf{a}_\rho$; (*c*) in spherical coordinates at $P_C(2, 30°, 20°)$ if $\mathbf{H} = \frac{1}{\sin\theta}\mathbf{a}_\theta$.

Ans. $-16\mathbf{a}_x + 9\mathbf{a}_y + 16\mathbf{a}_z$ A/m^2; $0.0549\mathbf{a}_z$ A/m^2; $\mathbf{a}_\phi$ A/m^2

8.4 STOKES' THEOREM

Although the last section was devoted primarily to a discussion of the curl opeation, the contribution to the subject of magnetic fields should not be overlooked. From Ampère's circuital law we derived one of Maxwell's equations, $\nabla \times \mathbf{H} = \mathbf{J}$. This latter equation should be considered the point form of Ampère's circuital law and applies on a "per-unit-area" basis. In this section we shall again devote a major share of the material to the mathematical theorem known as Stokes' theorem, but in the process we shall show that we may obtain Ampère's circuital law from $\nabla \times \mathbf{H} = \mathbf{J}$. In other words, we are then prepared to

obtain the integral form from the point form or to obtain the point form from the integral form.

Consider the surface S of Fig. 8.16 which is broken up into incremental surfaces of area ΔS. If we apply the definition of the curl to one of these incremental surfaces, then

$$\frac{\oint \mathbf{H} \cdot d\mathbf{L}_{\Delta S}}{\Delta S} \doteq (\nabla \times \mathbf{H})_N$$

where the N subscript again indicates the right-hand normal to the surface. The subscript on $d\mathbf{L}_{\Delta S}$ indicates that the closed path is the perimeter of an incremental area ΔS. This result may also be written

$$\frac{\oint \mathbf{H} \cdot d\mathbf{L}_{\Delta S}}{\Delta S} \doteq (\nabla \times \mathbf{H}) \cdot \mathbf{a}_N$$

or

$$\oint \mathbf{H} \cdot d\mathbf{L}_{\Delta S} \doteq (\nabla \times \mathbf{H} \cdot \mathbf{a}_N \Delta S = (\nabla \times \mathbf{H}) \cdot \Delta \mathbf{S}$$

where $\mathbf{a}_N$ is a unit vector in the direction of the right-hand normal to ΔS.

Now let us determine this circulation for every ΔS comprising S and sum the results. As we evaluate the closed line integral for each ΔS, some cancellation will occur because every *interior* wall is covered once in each direction. The only boundaries on which cancellation cannot occur form the outside boundary, the path enclosing S. Therefore we have

$$\boxed{\oint \mathbf{H} \cdot d\mathbf{L} \equiv \int_S (\nabla \times \mathbf{H}) \cdot d\mathbf{S}} \tag{30}$$

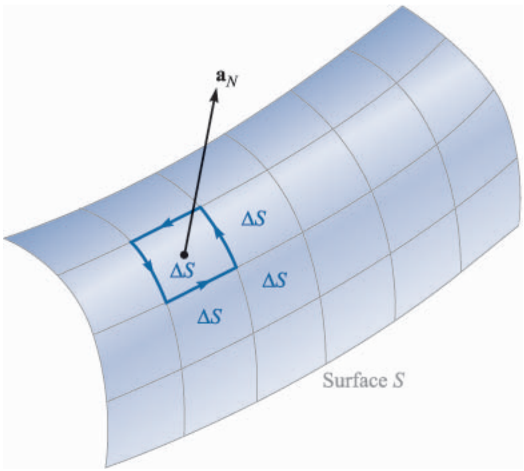


FIGURE 8.16
The sum of the closed line integrals about the perimeter of every ΔS is the same as the closed line integral about the perimeter of S because of cancellation on every interior path.

where $d\mathbf{L}$ is taken only on the perimeter of S.

Equation (30) is an identity, holding for any vector field, and is known as *Stokes' theorem.*

Example 8.3

A numerical example may help to illustrate the geometry involved in Stokes' theorem. Consider the portion of a sphere shown in Fig. 8.17. The surface is specified by $r = 4$, $0 \le \theta \le 0.1\pi$, $0 \le \phi \le 0.3\pi$, and the closed path forming its perimeter is composed of three circular arcs. We are given the field $\mathbf{H} = 6r\sin\phi\mathbf{a}_r + 18r\sin\theta\cos\phi\mathbf{a}_\phi$ and are asked to evaluate each side of Stokes' theorem.

Solution. The first path segment is described in spherical coordinates by $r = 4$, $0 \le \theta \le 0.1\pi$, $\phi = 0$; the second one by $r = 4$, $\theta = 0.1\pi$, $0 \le \phi \le 0.3\pi$; and the third by $r = 4$, $0 \le \theta \le 0.1\pi$, $\phi = 0.3\pi$. The differential path element $d\mathbf{L}$ is the vector sum of the three differential lengths of the spherical coordinate system first discussed in Sec. 1.9,

$$d\mathbf{L} = dr\,\mathbf{a}_r + r\,d\theta\,\mathbf{a}_\theta + r\sin\theta\,d\phi\,\mathbf{a}_\phi$$

The first term is zero on all three segments of the path since $r = 4$ and $dr = 0$, the second is zero on segment 2 since θ is constant, and the third term is zero on both segments 1 and 3. Thus

$$\oint \mathbf{H}\cdot d\mathbf{L} = \int_1 H_\theta r\,d\theta + \int_2 H_\phi r\sin\theta\,d\phi + \int_3 H_\theta r\,d\theta$$

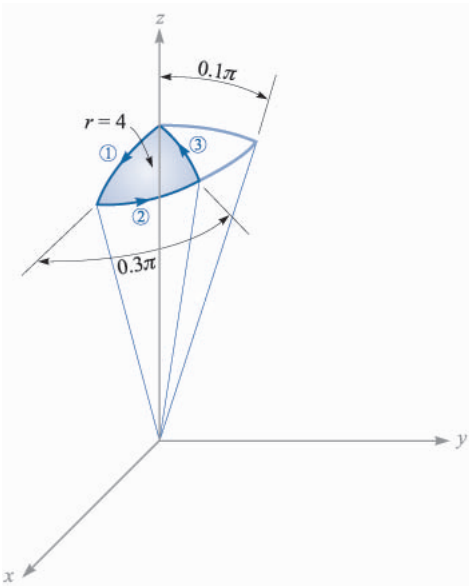


FIGURE 8.17
A portion of a spherical cap is used as a surface and a closed path to illustrate Stokes' theorem.

Since $H_\theta = 0$, we have only the second integral to evaluate,

$$\oint \mathbf{H} \cdot d\mathbf{L} = \int_0^{0.3\pi} [18(4) \sin 0.1\pi \cos\phi] 4 \sin 0.1\pi d\phi$$
$$= 288 \sin^2 0.1\pi \sin 0.3\pi = 22.2 \text{ A}$$

We next attack the surface integral. First, we use (26) to find

$$\nabla \times \mathbf{H} = \frac{1}{r \sin\theta}(36r \sin\theta \cos\theta \cos\phi)\mathbf{a}_r + \frac{1}{r}\left(\frac{1}{\sin\theta} 6r \cos\phi - 36r \sin\theta \cos\phi\right)\mathbf{a}_\theta$$

Since $d\mathbf{S} = r^2 \sin\theta \, d\theta \, d\phi \, \mathbf{a}_r$, the integral is

$$\int_S (\nabla \times \mathbf{H}) \cdot d\mathbf{S} = \int_0^{0.3\pi} \int_0^{0.1\pi} (36 \cos\theta \cos\phi) 16 \sin\theta \, d\theta \, d\phi$$
$$= \int_0^{0.3\pi} 576(\tfrac{1}{2} \sin^2\theta)\Big|_0^{0.1\pi} \cos\phi \, d\phi$$
$$= 288 \sin^2 0.1\pi \sin 0.3\pi = 22.2 \text{ A}$$

Thus, the results check Stokes' theorem, and we note in passing that a current of 22.2 A is flowing upward through this section of a spherical cap.

Next, let us see how easy it is to obtain Ampère's circuital law from $\nabla \times \mathbf{H} = \mathbf{J}$. We merely have to dot each side by $d\mathbf{S}$, integrate each side over the same (open) surface S, and apply Stokes' theorem:

$$\int_S (\nabla \times \mathbf{H}) \cdot d\mathbf{S} = \int_S \mathbf{J} \cdot d\mathbf{S} = \oint \mathbf{H} \cdot d\mathbf{L}$$

The integral of the current density over the surface S is the total current I passing through the surface, and therefore

$$\oint \mathbf{H} \cdot d\mathbf{L} = I$$

This short derivation shows clearly that the current I, described as being "enclosed by the closed path," is also the current passing through any of the infinite number of surfaces which have the closed path as a perimeter.

Stokes' theorem relates a surface integral to a closed line integral. It should be recalled that the divergence theorem relates a volume integral to a closed surface integral. Both theorems find their greatest use in general vector proofs. As an example, let us find another expression for $\nabla \cdot \nabla \times \mathbf{A}$, where $\mathbf{A}$ represents any vector field. The result must be a scalar (why?), and we may let this scalar be T, or

$$\nabla \cdot \nabla \times \mathbf{A} = T$$

Multiplying by dv and integrating throughout any volume v,

$$\int_{\text{vol}} (\nabla \cdot \nabla \times \mathbf{A}) dv = \int_{\text{vol}} T \, dv$$

we first apply the divergence theorem to the left side, obtaining

$$\oint_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = \int_{\text{vol}} T\, dv$$

The left side is the surface integral of the curl of **A** over the *closed* surface surrounding the volume v. Stokes' theorem relates the surface integral of the curl of **A** over the *open* surface enclosed by a given closed path. If we think of the path as the opening of a laundry bag and the open surface as the surface of the bag itself, we see that as we gradually approach a closed surface by pulling on the drawstrings, the closed path becomes smaller and smaller and finally disappears as the surface becomes closed. Hence the application of Stokes' theorem to a *closed* surface produces a zero result, and we have

$$\int_{\text{vol}} T\, dv = 0$$

Since this is true for any volume, it is true for the differential volume dv,

$$T\, dv = 0$$

and therefore

$$T = 0$$

or

$$\boxed{\nabla \cdot \nabla \times \mathbf{A} \equiv 0} \tag{31}$$

Equation (31) is a useful identity of vector calculus.[6] Of course, it may also be proven easily by direct expansion in cartesian coordinates.

Let us apply the identity to the non-time-varying magnetic field for which

$$\nabla \times \mathbf{H} = \mathbf{J}$$

This shows quickly that

$$\mathbf{V} \cdot \mathbf{J} = 0$$

which is the same result we obtained earlier in the chapter by using the continuity equation.

Before introducing several new magnetic field quantities in the following section, we may review our accomplishments at this point. We initially accepted the Biot-Savart law as an experimental result,

$$\mathbf{H} = \oint \frac{I\, d\mathbf{L} \times \mathbf{a}_R}{4\pi R^2}$$

---

[6] This and other vector identities are tabulated in Appendix A.3.

and tentatively accepted Ampère's circuital law, subject to later proof,

$$\oint \mathbf{H} \cdot d\mathbf{L} = I$$

From Ampère's circuital law the definition of curl led to the point form of this same law,

$$\nabla \times \mathbf{H} = \mathbf{J}$$

We now see that Stokes' theorem enables us to obtain the integral form of Ampère's circuital law from the point of form.

✔ **D8.6.** Evaluate both sides of Stokes' theorem for the field $\mathbf{H} = 6xy\mathbf{a}_x - 3y^2\mathbf{a}_y$ A/m and the rectangular path around the region, $2 \le x \le 5$, $-1 \le y \le 1$, $z = 0$. Let the positive direction of $d\mathbf{S}$ be $\mathbf{a}_z$.

Ans. -126 A; -126 A

8.5 MAGNETIC FLUX AND MAGNETIC FLUX DENSITY

In free space, let us define the *magnetic flux density* **B** as

$$\boxed{\mathbf{B} = \mu_0 \mathbf{H}} \qquad \text{(free space only)} \tag{32}$$

where **B** is measured in webers per square meter (Wb/m^2) or in a newer unit adopted in the International System of Units, tesla (T). An older unit that is often used for magnetic flux density is the gauss (G), where 1 T or 1 Wb/m^2 is the same as 10 000 G. The constant μ_0 is not dimensionless and has the *defined value* for free space, in henrys per meter (H/m), of

$$\boxed{\mu_0 = 4\pi \times 10^{-7} \text{ H/m}} \tag{33}$$

The name given to μ_0 is the *permeability* of free space.

We should note that since **H** is measured in amperes per meter, the weber is dimensionally equal to the product of henrys and amperes. Considering the henry as a new unit, the weber is merely a convenient abbreviation for the product of henrys and amperes. When time-varying fields are introduced, it will be shown that a weber is also equivalent to the product of volts and seconds.

The magnetic-flux-density vector **B**, as the name weber per square meter implies, is a member of the flux-density family of vector fields. One of the possible analogies between electric and magnetic fields[7] compares the laws of Biot-Savart and Coulomb, thus establishing an analogy between **H** and **E**. The

[7] An alternate analogy is presented in Sec. 10.2.

relations $\mathbf{B} = \mu_0 \mathbf{H}$ and $\mathbf{D} = \epsilon_0 \mathbf{E}$ then lead to an analogy between $\mathbf{B}$ and $\mathbf{D}$. If $\mathbf{B}$ is measured in teslas or webers per square meter, then magnetic flux should be measured in webers. Let us represent magnetic flux by Φ and define Φ as the flux passing through any designated area,

$$\boxed{\Phi = \int_S \mathbf{B} \cdot d\mathbf{S} \text{ Wb}} \tag{34}$$

Our analogy should now remind us of the electric flux Ψ, measured in coulombs, and of Gauss's law, which states that the total flux passing through any closed surface is equal to the charge enclosed,

$$\Psi = \oint_S \mathbf{D} \cdot d\mathbf{S} = Q$$

The charge Q is the source of the lines of electric flux and these lines begin and terminate on positive and negative charge, respectively.

No such source has ever been discovered for the lines of magnetic flux. In the example of the infinitely long straight filament carrying a direct current I, the $\mathbf{H}$ field formed concentric circles about the filament. Since $\mathbf{B} = \mu_0 \mathbf{H}$, the $\mathbf{B}$ field is of the same form. The magnetic flux lines are closed and do not terminate on a "magnetic charge." For this reason Gauss's law for the magnetic field is

$$\boxed{\oint_S \mathbf{B} \cdot d\mathbf{S} = 0} \tag{35}$$

and application of the divergence theorem shows us that

$$\boxed{\nabla \cdot \mathbf{B} = 0} \tag{36}$$

We have not proved (35) or (36) but have only suggested the truth of these statements by considering the single field of the infinite filament. It is possible to show that (35) or (36) follows from the Biot-Savart law and the definition of $\mathbf{B}$, $\mathbf{B} = \mu_0 \mathbf{H}$, but this is another proof which we shall postpone to Sec. 8.7.

Equation (36) is the last of Maxwell's four equations as they apply to static electric fields and steady magnetic fields. Collecting these equations, we then have for static electric fields and steady magnetic fields

$$\boxed{\begin{aligned} \nabla \cdot \mathbf{D} &= \rho_v \\ \nabla \times \mathbf{E} &= 0 \\ \nabla \times \mathbf{H} &= \mathbf{J} \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned}} \tag{37}$$

To these equations we may add the two expressions relating **D** to **E** and **B** to **H** in free space,

$$\mathbf{D} = \epsilon_0 \mathbf{E} \tag{38}$$

$$\mathbf{B} = \mu_0 \mathbf{H} \tag{39}$$

We have also found it helpful to define an electrostatic potential,

$$\mathbf{E} = -\nabla V \tag{40}$$

and we shall discuss a potential for the steady magnetic field in the following section. In addition, we have extended our coverage of electric fields to include conducting materials and dielectrics, and we have introduced the polarization **P**. A similar treatment will be applied to magnetic fields in the next chapter.

Returning to (37), it may be noted that these four equations specify the divergence and curl of an electric and a magnetic field. The corresponding set of four integral equations that apply to static electric fields and steady magnetic fields is

$$\begin{aligned}
&\oint_S \mathbf{D} \cdot d\mathbf{S} = Q = \int_{\text{vol}} \rho_v dv \\
&\oint \mathbf{E} \cdot d\mathbf{L} = 0 \\
&\oint \mathbf{H} \cdot d\mathbf{L} = I = \int_S \mathbf{J} \cdot d\mathbf{S} \\
&\oint_S \mathbf{B} \cdot d\mathbf{S} = 0
\end{aligned} \tag{41}$$

Our study of electric and magnetic fields would have been much simpler if we could have begun with either set of equations, (37) or (41). With a good knowledge of vector analysis, such as we should now have, either set may be readily obtained from the other by applying the divergence theorem of Stokes' theorem. The various experimental laws can be obtained easily from these equations.

As an example of the use of flux and flux density in magnetic fields, let us find the flux between the conductors of the coaxial line of Fig. 8.8*a*. The magnetic field intensity was found to be

$$H_\phi = \frac{I}{2\pi\rho} \quad (a < \rho < b)$$

and therefore

$$\mathbf{B} = \mu_0 \mathbf{H} = \frac{\mu_0 I}{2\pi\rho} \mathbf{a}_\phi$$

The magnetic flux contained between the conductors in a length d is the flux crossing any radial plane extending from $\rho = a$ to $\rho = b$ and from, say, $z = 0$ to $z = d$

$$\Phi = \int_S \mathbf{B} \cdot d\mathbf{S} = \int_0^d \int_a^b \frac{\mu_0 I}{2\pi\rho} \mathbf{a}_\phi \cdot d\rho\, dz\, \mathbf{a}_\phi$$

or

$$\Phi = \frac{\mu_0 I d}{2\pi} \ln \frac{b}{a} \tag{42}$$

This expression will be used later to obtain the inductance of the coaxial transmission line.

✔ **D8.7.** A solid conductor of circular cross section is made of a homogeneous nonmagnetic material. If the radius $a = 1$ mm, the conductor axis lies on the z axis, and the total current in the $\mathbf{a}_z$ direction is 20 A, find: (*a*) H_ϕ at $\rho = 0.5$ mm; (*b*) B_ϕ at $\rho = 0.8$ mm; (*c*) the total magnetic flux per unit length inside the conductor; (*d*) the total flux for $\rho < 0.5$ mm; (*e*) the total magnetic flux outside the conductor.

Ans. 1592 A/m; 3.2 mT; 2μWb; 0.5 μWb; ∞

8.6 THE SCALAR AND VECTOR MAGNETIC POTENTIALS

The solution of electrostatic field problems is greatly simplified by the use of the scalar electrostatic potential V. Although this potential possesses a very real physical significance for us, it is mathematically no more than a stepping-stone which allows us to solve a problem by several smaller steps. Given a charge configuration, we may first find the potential and then from it the electric field intensity.

We should question whether or not such assistance is available in magnetic fields. Can we define a potential function which may be found from the current distribution and from which the magnetic fields may be easily determined? Can a scalar magnetic potential be defined, similar to the scalar electrostatic potential? We shall show in the next few pages that the answer to the first question is "yes," but the second must be answered "sometimes." Let us attack the last question first by assuming the existence of a scalar magnetic potential, which we designate V_m, whose negative gradient gives the magnetic field intensity

$$\mathbf{H} = -\nabla V_m$$

The selection of the negative gradient will provide us with a closer analogy to the electric potential and to problems which we have already solved.

This definition must not conflict with our previous results for the magnetic field, and therefore

$$\nabla \times \mathbf{H} = \mathbf{J} = \nabla \times (-\nabla V_m)$$

However, the curl of the gradient of any scalar is identically zero, a vector identity the proof of which is left for a leisure moment. Therefore we see that if **H** is to be defined as the gradient of a scalar magnetic potential, then current density must be zero throughout the region in which the scalar magnetic potential is so defined. We then have

$$\boxed{\mathbf{H} = -\nabla V_m \quad (\mathbf{J} = 0)} \tag{43}$$

Since many magnetic problems involve geometries in which the current-carrying conductors occupy a relatively small fraction of the total region of interest, it is evident that a scalar magnetic potential can be useful. The scalar magnetic potential is also applicable in the case of permanent magnets. The dimensions of V_m are obviously amperes.

This scalar potential also satisfies Laplace's equation. In free space,

$$\nabla \cdot \mathbf{B} = \mu_0 \nabla \cdot \mathbf{H} = 0$$

and hence

$$\mu_0 \nabla \cdot (-\nabla V_m) = 0$$

or

$$\boxed{\nabla^2 V_m = 0 \quad (\mathbf{J} = 0)} \tag{44}$$

We shall see later that V_m continues to satisfy Laplace's equation in homogeneous magnetic materials; it is not defined in any region in which current density is present.

Although we shall consider the scalar magnetic potential to a much greater extent in the next chapter, when we introduce magnetic materials and discuss the magnetic circuit, one difference between V and V_m should be pointed out now: V_m is not a single-valued function of position. The electric potential V is single-valued; once a zero reference is assigned, there is only one value of V associated with each point in space. Such is not the case with V_m. Consider the cross section of the coaxial line shown in Fig. 8.18. In the region $a < \rho < b$, $\mathbf{J} = 0$, and we may establish a scalar magnetic potential. The value of **H** is

$$\mathbf{H} = \frac{I}{2\pi\rho}\mathbf{a}_\phi$$

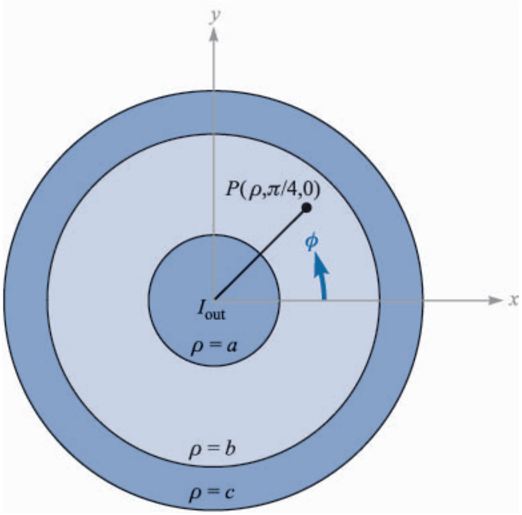


FIGURE 8.18
The scalar magnetic potential V_m is a multi-valued function of ϕ in the region $a < \rho < b$. The electrostatic potential is always single-valued.

where I is the total current flowing in the $\mathbf{a}_z$ direction in the inner conductor. Let us find V_m by integrating the appropriate component of the gradient. Applying (43),

$$\frac{I}{2\pi\rho} = -\nabla V_m\Big|_\phi = -\frac{1}{\rho}\frac{\partial V_m}{\partial\phi}$$

or

$$\frac{\partial V_m}{\partial\phi} = -\frac{I}{2\pi}$$

Thus

$$V_m = -\frac{I}{2\pi}\phi$$

where the constant of integration has been set equal to zero. What value of potential do we associate with point P, where $\phi = \pi/4$? If we let V_m be zero at $\phi = 0$ and proceed counterclockwise around the circle, the magnetic potential goes negative linearly. When we have made one circuit, the potential is $-I$, but that was the point at which we said the potential was zero a moment ago. At P, then, $\phi = \pi/4,\ 9\pi/4,\ 17\pi/4, \ldots$, or $-7\pi/4,\ -15\pi/4,\ -23\pi/4, \ldots$, or

$$V_{mP} = \frac{I}{2\pi}(2n - \tfrac{1}{4})\pi \quad (n = 0, \pm 1, \pm 2, \ldots)$$

or

$$V_{mP} = I(n - \tfrac{1}{8}) \quad (n = 0, \pm 1, \pm 2, \ldots)$$

The reason for this multivaluedness may be shown by a comparison with the electrostatic case. There, we know that

$$\nabla \times \mathbf{E} = 0$$

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0$$

and therefore the line integral

$$V_{ab} = -\int_b^a \mathbf{E} \cdot d\mathbf{L}$$

is independent of the path. In the magnetostatic case, however,

$$\nabla \times \mathbf{H} = 0 \quad (\text{wherever } \mathbf{J} = 0)$$

but

$$\oint \mathbf{H} \cdot d\mathbf{L} = I$$

even if **J** is zero along the path of integration. Every time we make another complete lap around the current, the result of the integration increases by I. If no current I is enclosed by the path, then a single-valued potential function may be defined. In general, however,

$$V_{m,ab} = -\int_b^a \mathbf{H} \cdot d\mathbf{L} \quad (\text{specified path}) \tag{45}$$

where a specific path or type of path must be selected. We should remember that the electrostatic potential V is a conservative field; the magnetic scalar potential V_m is not a conservative field. In our coaxial problem let us erect a barrier[8] at $\phi = \pi$; we agree not to select a path which crosses this plane. Therefore we cannot encircle I, and a single-valued potential is possible. The result is seen to be

$$V_m = -\frac{I}{2\pi}\phi \quad (-\pi < \phi < \pi)$$

and

$$V_{mP} = -\frac{I}{8} \quad \left(\phi = \frac{\pi}{4}\right)$$

The scalar magnetic potential is evidently the quantity whose equipotential surfaces will form curvilinear squares with the streamlines of **H** in Fig. 8.4. This is one more facet of the analogy between electric and magnetic fields about which we will have more to say in the next chapter.

---

[8] This corresponds to the more precise mathematical term "branch cut."

Let us temporarily leave the scalar magnetic potential now and investigate a vector magnetic potential. This vector field is one which is extremely useful in studying radiation from antennas, from apertures, and radiation leakage from transmission lines, waveguides, and microwave ovens. The vector magnetic potential may be used in regions where the current density is zero or nonzero, and we shall also be able to extend it to the time-varying case later.

Our choice of a vector magnetic potential is indicated by noting that

$$\nabla \cdot \mathbf{B} = 0$$

Next, a vector identity which we proved in Sec. 8.4 shows that the divergence of the curl of any vector field is zero. Therefore we select

$$\boxed{\mathbf{B} = \nabla \times \mathbf{A}} \tag{46}$$

where **A** signifies a *vector magnetic potential*, and we automatically satisfy the condition that the magnetic flux density shall have zero divergence. The **H** field is

$$\mathbf{H} = \frac{1}{\mu_0} \nabla \times \mathbf{A}$$

and

$$\nabla \times \mathbf{H} = \mathbf{J} = \frac{1}{\mu_0} \nabla \times \nabla \times \mathbf{A}$$

The curl of the curl of a vector field is not zero and is given by a fairly complicated expression,[9] which we need not know now in general form. In specific cases for which the form of **A** is known, the curl operation may be applied twice to determine the current density.

Equation (46) serves as a useful definition of the *vector magnetic potential* **A**. Since the curl operation implies differentiation with respect to a length, the units of **A** are webers per meter.

As yet we have seen only that the definition for **A** does not conflict with any previous results. It still remains to show that this particular definition can help us to determine magnetic fields more easily. We certainly cannot identify **A** with any easily measured quantity or history-making experiment.

We shall show in the following section that, given the Biot-Savart law, the definition of **B**, and the definition of **A**, then **A** may be determined from the differential current elements by

---

[9] $\nabla \times \nabla \times \mathrm{A} \equiv \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathrm{A}$. In cartesian coordinates, it may be shown that $\nabla^2 \mathrm{A} \equiv \nabla^2 A_x \mathbf{a}_x + \nabla^2 A_y \mathbf{a}_y + \nabla^2 A_z \mathbf{a}_z$. In other coordinate systems, $\nabla^2 \mathrm{A}$ may be found by evaluating the second-order partial derivatives in $\nabla^2 \mathrm{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla \times \nabla \times \mathrm{A}$.

$$\boxed{\mathbf{A} = \oint \frac{\mu_0 I \, d\mathbf{L}}{4\pi R}} \tag{47}$$

The significance of the terms in (47) is the same as in the Biot-Savart law; a direct current I flows along a filamentary conductor of which any differential length $d\mathbf{L}$ is distant R from the point at which $\mathbf{A}$ is to be found. Since we have defined $\mathbf{A}$ only through specification of its curl, it is possible to add the gradient of any scalar field to (47) without changing $\mathbf{B}$ or $\mathbf{H}$, for the curl of the gradient is identically zero. In steady magnetic fields, it is customary to set this possible added term equal to zero.

The fact that $\mathbf{A}$ is a vector magnetic *potential* is more apparent when (47) is compared with the similar expression for the electrostatic potential,

$$V = \int \frac{\rho_L dL}{4\pi\epsilon_0 R}$$

Each expression is the integral along a line source, in one case line charge and in the other case line current; each integrand is inversely proportional to the distance from the source to the point of interest; and each involves a characteristic of the medium (here free space), the permeability or the permittivity.

Equation (47) may be written in differential form,

$$d\mathbf{A} = \frac{\mu_0 I \, d\mathbf{L}}{4\pi R} \tag{48}$$

if we again agree not to attribute any physical significance to any magnetic fields we obtain from (48) until the *entire closed path in which the current flows is considered.*

With this reservation, let us go right ahead and consider the vector magnetic potential field about a differential filament. We locate the filament at the origin in free space, as shown in Fig. 8.19, and allow it to extend in the positive z direction so that $d\mathbf{L} = dz\,\mathbf{a}_z$. We use cylindrical coordinates to find $d\mathbf{A}$ at the point (ρ, ϕ, z):

$$d\mathbf{A} = \frac{\mu_0 I \, dz\,\mathbf{a}_z}{4\pi\sqrt{\rho^2 + z^2}}$$

or

$$d\mathbf{A}_z = \frac{\mu_0 I \, dz}{4\pi\sqrt{\rho^2 + z^2}} \qquad dA_\phi = 0 \quad dA_\rho = 0 \tag{49}$$

We note that the direction of $d\mathbf{A}$ is the same as that of $I\,d\mathbf{L}$. Each small section of a current-carrying conductor produces a contribution to the total vector magnetic potential which is in the same direction as the current flow in the conductor. The magnitude of the vector magnetic potential varies inversely as the distance to the current element, being strongest in the neighborhood of the

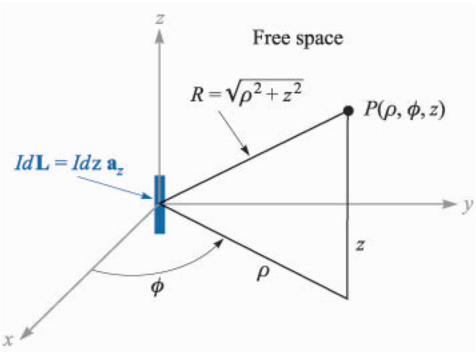


FIGURE 8.19
The differential current element $I\, dz\mathbf{a}_z$ at the origin establishes the differential vector magnetic potential field, $d\mathbf{A} = \dfrac{\mu_0 I\, dz\mathbf{a}_z}{4\pi\sqrt{\rho^2+z^2}}$ at $P(\rho, \phi, z)$.

current and gradually falling off to zero at distant points. Skilling[10] describes the vector magnetic potential field as "like the current distribution but fuzzy around the edges, or like a picture of the current out of focus."

In order to find the magnetic field intensity, we must take the curl of (49) in cylindrical coordinates, leading to

$$d\mathbf{H} = \frac{1}{\mu_0}\nabla \times d\mathbf{A} = \frac{1}{\mu_0}\left(-\frac{\partial dA_z}{\partial \rho}\right)\mathbf{a}_\phi$$

or

$$d\mathbf{H} = \frac{I\, dz}{4\pi}\frac{\rho}{(\rho^2+z^2)^{3/2}}\mathbf{a}_\phi$$

which is easily shown to be the same as the value given by the Biot-Savart law.

Expressions for the vector magnetic potential $\mathbf{A}$ can also be obtained for a current source which is distributed. For a current sheet $\mathbf{K}$, the differential current element becomes

$$I\, d\mathbf{L} = \mathbf{K}\, dS$$

In the case of current flow throughout a volume with a density $\mathbf{J}$, we have

$$I\, d\mathbf{L} = \mathbf{J}\, dv$$

In each of these two expressions the vector character is given to the current. For the filamentary element it is customary, although not necessary, to use $I\, d\mathbf{L}$ instead of $\mathbf{I}\, dL$. Since the magnitude of the filamentary element is constant, we have chosen the form which allows us to remove one quantity from the integral. The alternative expressions for $\mathbf{A}$ are then

[10] See the Suggested References at the end of the chapter.

$$\mathbf{A} = \int_S \frac{\mu_0 \mathbf{K}\, dS}{4\pi R} \tag{50}$$

and

$$\mathbf{A} = \int_{\text{vol}} \frac{\mu_0 \mathbf{J}\, dv}{4\pi R} \tag{51}$$

Equations (47), (50), and (51) express the vector magnetic potential as an integration over all of its sources. From a comparison of the form of these integrals with those which yield the electrostatic potential, it is evident that once again the zero reference for **A** is at infinity, for no finite current element can produce any contribution as $R \to \infty$. We should remember that we very seldom used the similar expressions for V; too often our theoretical problems included charge distributions which extended to infinity and the result would be an infinite potential everywhere. Actually, we calculated very few potential fields until the differential form of the potential equation was obtained, $\nabla^2 V = -\rho_v/\epsilon$, or better yet, $\nabla^2 V = 0$. We were then at liberty to select our own zero reference.

The analogous expressions for **A** will be derived in the next section, and an example of the calculation of a vector magnetic potential field will be completed.

✔ **D8.8.** A current sheet, $\mathbf{K} = 2.4\mathbf{a}_z$ A/m, is present at the surface $\rho = 1.2$ in free space. (*a*) Find **H** for $\rho > 1.2$. Find V_m at $P(\rho = 1.5, \phi = 0.6\pi, z = 1)$ if: (*b*) $V_m = 0$ at $\phi = 0$ and there is a barrier at $\phi = \pi$; (*c*) $V_m = 0$ at $\phi = 0$ and there is a barrier at $\phi = \pi/2$; (*d*) $V_m = 0$ at $\phi = \pi$ and there is a barrier at $\phi = 0$; (*e*) $V_m = 5$ V at $\phi = \pi$ and there is a barrier at $\phi = 0.8\pi$.

Ans. $\dfrac{2.88}{\rho}\alpha_\phi$; -0.864 V; 2.02 V; 0.576 V; 2.70 V

✔ **D8.9.** The value of **A** within a solid nonmagnetic conductor of radius a carrying a total current I in the $\mathbf{a}_z$ direction may be found easily. Using the known value of **H** or **B** for $\rho < a$, then (46) may be solved for **A**. Select $A = (\mu_0 I \ln 5)/2\pi$ at $\rho = a$ (to correspond with an example in the next section) and find **A** at $\rho =$: (*a*) 0; (*b*) 0.25a; (*c*) 0.75a; (*d*) a.

Ans. 0.422 μ Wb/m; 0.416 μ Wb/m; 0.366 μ Wb/m; 0.322 μ Wb/m

8.7 DERIVATION OF THE STEADY-MAGNETIC-FIELD LAWS

We shall now carry out our threat to supply the promised proofs of the several relationships between the magnetic field quantities. All these relationships may be obtained from the definitions of **H**,

$$\mathbf{H} = \oint \frac{I\, d\mathbf{L} \times \mathbf{a}_R}{4\pi R^2} \tag{3}$$

of **B** (in free space),

$$\mathbf{B} = \mu_0 \mathbf{H} \tag{32}$$

and of **A**,

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{46}$$

Let us first assume that we may express **A** by the last equation of the preceding section,

$$\mathbf{A} = \int_{\text{vol}} \frac{\mu_0 \mathbf{J}\, dv}{4\pi R} \tag{51}$$

and then demonstrate the correctness of (51) by showing that (3) follows. First we should add subscripts to indicate the point at which the current element is located (x_1, y_1, z_1) and the point at which **A** is given (x_2, y_2, z_2). The differential volume element dv is then written dv_1 and in cartesian coordinates would be $dx_1\, dy_1\, dz_1$. The variables of integration are x_1, y_1, and z_1. Using these subscripts, then,

$$\mathbf{A}_2 = \int_{\text{vol}} \frac{\mu_0 \mathbf{J}_1 dv_1}{4\pi R_{12}} \tag{52}$$

From (32) and (46) we have

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} = \frac{\nabla \times \mathbf{A}}{\mu_0} \tag{53}$$

To show that (3) follows from (52), it is necessary to substitute (52) into (53). This step involves taking the curl of $\mathbf{A}_2$, a quantity expressed in terms of the variables x_2, y_2, and z_2, and the curl therefore involves partial derivatives with respect to x_2, y_2, and z_2. We do this, placing a subscript on the del operator to remind us of the variables involved in the partial differentiation process,

$$\mathbf{H}_2 = \frac{\nabla_2 \times \mathbf{A}_2}{\mu_0} = \frac{1}{\mu_0} \nabla_2 \times \int_{\text{vol}} \frac{\mu_0 \mathbf{J}_1 dv_1}{4\pi R_{12}}$$

The order of partial differentiation and integration is immaterial, and $\mu_0/4\pi$ is constant, allowing us to write

$$\mathbf{H}_2 = \frac{1}{4\pi} \int_{\text{vol}} \nabla_2 \times \frac{\mathbf{J}_1 dv_1}{R_{12}}$$

The curl operation within the integrand represents partial differentiation with respect to x_2, y_2, and z_2. The differential volume element dv_1 is a scalar and a function only of x_1, y_1, and z_1. Consequently, it may be factored out of the curl operation as any other constant, leaving

$$\mathbf{H}_2 = \frac{1}{4\pi}\int_{\text{vol}}\left(\nabla_2 \times \frac{\mathbf{J}_1}{R_{12}}\right)dv_1 \tag{54}$$

The curl of the product of a scalar and a vector is given by an identity which may be checked by expansion in cartesian coordinates or gratefully accepted from Appendix A.3,

$$\nabla \times (S\mathbf{V}) \equiv (\nabla S) \times \mathbf{V} + S(\nabla \times \mathbf{V}) \tag{55}$$

This identity is used to expand the integrand of (54),

$$\mathbf{H}_2 = \frac{1}{4\pi}\int_{\text{vol}}\left[\left(\nabla_2 \frac{1}{R_{12}}\right) \times \mathbf{J}_1 + \frac{1}{R_{12}}(\nabla_2 \times \mathbf{J}_1)\right]dv_1 \tag{56}$$

The second term of this integrand is zero, because $\nabla \times \mathbf{J}_1$ indicates partial derivatives of a function of x_1, y_1, and z_1, taken with respect to the variables x_2, y_2, and z_2; the first set of variables is not a function of the second set, and all partial derivatives are zero.

The first term of the integrand may be determined by expressing R_{12} in terms of the coordinate values,

$$R_{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

and taking the gradient of its reciprocal. Problem 42 shows that the result is

$$\nabla_2 \frac{1}{R_{12}} = -\frac{\mathbf{R}_{12}}{R_{12}^3} = -\frac{\mathbf{a}_{R12}}{R_{12}^2}$$

Substituting this result into (56), we have

$$\mathbf{H}_2 = -\frac{1}{4\pi}\int_{\text{vol}} \frac{\mathbf{a}_{R12} \times \mathbf{J}_1}{R_{12}^2}\, dv_1$$

or

$$\mathbf{H}_2 = \int_{\text{vol}} \frac{\mathbf{J}_1 \times \mathbf{a}_{R12}}{4\pi R_{12}^2}\, dv_1$$

which is the equivalent of (3) in terms of current density. Replacing $\mathbf{J}_1\, dv_1$ by $I_1\, d\mathbf{L}_1$, we may rewrite the volume integral as a closed line integral,

$$\mathbf{H}_2 = \oint \frac{I_1 d\mathbf{L}_1 \times \mathbf{a}_{R12}}{4\pi R_{12}^2}$$

Equation (51) is therefore correct and agrees with the three definitions (3), (32), and (46).

Next we shall continue with our mathematical orgy and prove Ampère's circuital law in point form,

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{28}$$

Combining (28), (32), and (46), we obtain

$$\nabla \times \mathbf{H} = \nabla \times \frac{\mathbf{B}}{\mu_0} = \frac{1}{\mu_0} \nabla \times \nabla \times \mathbf{A} \tag{57}$$

We now need the expansion in cartesian coordinates for $\nabla \times \nabla \times \mathbf{A}$. Performing the indicated partial differentiations and collecting the resulting terms, we may write the result as

$$\boxed{\nabla \times \nabla \times \mathbf{A} \equiv \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}} \tag{58}$$

where

$$\boxed{\nabla^2 \mathbf{A} \equiv \nabla^2 A_x \mathbf{a}_x + \nabla^2 A_y \mathbf{a}_y + \nabla^2 A_z \mathbf{a}_z} \tag{59}$$

Equation (59) is the definition (in cartesian coordinates) of the *Laplacian of a vector*.

Substituting (58) into (57), we have

$$\nabla \times \mathbf{H} = \frac{1}{\mu_0}[\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}] \tag{60}$$

and now require expressions for the divergence and the Laplacian of $\mathbf{A}$.

We may find the divergence of $\mathbf{A}$ by applying the divergence operation to (52),

$$\nabla_2 \cdot \mathbf{A}_2 = \frac{\mu_0}{4\pi} \int_{\text{vol}} \nabla_2 \cdot \frac{\mathbf{J}_1}{R_{12}} dv_1 \tag{61}$$

and using the vector identity (44) of Sec. 4.8,

$$\nabla \cdot (S\mathbf{V}) \equiv \mathbf{V} \cdot (\nabla S) + S(\nabla \cdot \mathbf{V})$$

Thus,

$$\nabla_2 \cdot \mathbf{A}_2 = \frac{\mu_0}{4\pi} \int_{\text{vol}} \left[\mathbf{J}_1 \cdot \left(\nabla_2 \frac{1}{R_{12}} \right) + \frac{1}{R_{12}} (\nabla_2 \cdot \mathbf{J}_1) \right] dv_1 \tag{62}$$

The second part of the integrand is zero, because $\mathbf{J}_1$ is not a function of x_2, y_2, and z_2.

We have already used the result that $\nabla_2(1/R_{12}) = -\mathbf{R}_{12}/R_{12}^3$, and it is just as easily shown that

$$\nabla_1 \frac{1}{R_{12}} = \frac{\mathbf{R}_{12}}{R_{12}^3}$$

or that

$$\nabla_1 \frac{1}{R_{12}} = -\nabla_2 \frac{1}{R_{12}}$$

Equation (62) can therefore be written as

$$\nabla_2 \cdot \mathbf{A}_2 = \frac{\mu_0}{4\pi} \int_{\text{vol}} \left[-\mathbf{J}_1 \cdot \left(\nabla_1 \frac{1}{R_{12}} \right) \right] dv_1$$

and the vector identity applied again,

$$\nabla_2 \cdot \mathbf{A}_2 = \frac{\mu_0}{4\pi} \int_{\text{vol}} \left[\frac{1}{R_{12}} (\nabla_1 \cdot \mathbf{J}_1) - \nabla_1 \cdot \left(\frac{\mathbf{J}_1}{R_{12}} \right) \right] dv_1 \tag{63}$$

Since we are concerned only with steady magnetic fields, the continuity equation shows that the first term of (63) is zero. Application of the divergence theorem to the second term gives

$$\nabla_2 \cdot \mathbf{A}_2 = -\frac{\mu_0}{4\pi} \oint_{S_1} \frac{\mathbf{J}_1}{R_{12}} \cdot d\mathbf{S}_1$$

where the surface S_1 encloses the volume throughout which we are integrating. This volume must include all the current, for the original integral expression for **A** was an integration such as to include the effect of all the current. Since there is no current outside this volume (otherwise we should have had to increase the volume to include it), we may integrate over a slightly larger volume or a slightly larger enclosing surface without changing **A**. On this larger surface the current density $\mathbf{J}_1$ must be zero, and therefore the closed surface integral is zero, since the integrand is zero. Hence the divergence of **A** is zero.

In order to find the Laplacian of the vector **A** let us compare the x component of (51) with the similar expression for electrostatic potential,

$$A_x = \int_{\text{vol}} \frac{\mu_0 J_x dv}{4\pi R} \qquad V = \int_{\text{vol}} \frac{\rho_v \, dv}{4\pi\epsilon_0 R}$$

We note that one expression can be obtained from the other by a straightforward change of variable, J_x for ρ_v, μ_0 for $1/\epsilon_0$, and A_x for V. However, we have derived some additional information about the electrostatic potential which we shall not have to repeat now for the x component of the vector magnetic potential. This takes the form of Poisson's equation,

$$\nabla^2 V = -\frac{\rho_v}{\epsilon_0}$$

which becomes, after the change of variables,

$$\nabla^2 A_x = -\mu_0 J_x$$

Similarly, we have

$$\nabla^2 A_y = -\mu_0 J_y$$

and

$$\nabla^2 A_z = -\mu_0 J_z$$

or

$$\boxed{\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}} \tag{64}$$

Returning to (60), we can now substitute for the divergence and Laplacian of $\mathbf{A}$ and obtain the desired answer,

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{28}$$

We have already shown the use of Stokes' theorem in obtaining the integral form of Ampère's circuital law from (28) and need not repeat that labor here.

We thus have succeeded in showing that every result we have essentially pulled from thin air[11] for magnetic fields follows from the basic definitions of $\mathbf{H}$, $\mathbf{B}$, and $\mathbf{A}$. The derivations are not simple, but they should be understandable on a step-by-step basis. It is hoped that the procedure need never be committed to memory.

Finally, let us return to (64) and make use of this formidable second-order vector partial differential equation to find the vector magnetic potential in one simple example. We select the field between conductors of a coaxial cable, with radii of a and b as usual, and current I in the $\mathbf{a}_z$ direction in the inner conductor. Between the conductors, $\mathbf{J} = 0$, and therefore

$$\nabla^2 \mathbf{A} = 0$$

We have already been told (and Prob. 8.44 gives us the opportunity to check the results for ourselves) that the vector Laplacian may be expanded as the vector sum of the scalar Laplacians of the three components in cartesian coordinates,

$$\nabla^2 \mathbf{A} = \nabla^2 A_x \mathbf{a}_x + \nabla^2 A_y \mathbf{a}_y + \nabla^2 A_z \mathbf{a}_z$$

but such a relatively simple result is not possible in other coordinate systems. That is, in cylindrical coordinates, for example,

$$\nabla^2 \mathbf{A} \neq \nabla^2 A_\rho \mathbf{a}_\rho + \nabla^2 A_\phi \mathbf{a}_\phi + \nabla^2 A_z \mathbf{a}_z$$

However, it is not difficult to show for cylindrical coordinates that the z component of the vector Laplacian is the scalar Laplacian of the z component of $\mathbf{A}$, or

$$\nabla^2 \mathbf{A}\Big|_z = \nabla^2 A_z \tag{65}$$

and since the current is entirely in the z direction in this problem, $\mathbf{A}$ has only a z component. Therefore

$$\nabla^2 A_z = 0$$

[11] Free space.

or

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial A_z}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2 A_z}{\partial\phi^2}+\frac{\partial^2 A_z}{\partial z^2}=0$$

Thinking symmetrical thoughts about (51) shows us that A_z is a function only of ρ, and thus

$$\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{dA_z}{d\rho}\right)=0$$

We have solved this equation before, and the result is

$$A_z = C_1 \ln\rho + C_2$$

If we choose a zero reference at $\rho = b$, then

$$A_z = C_1 \ln\frac{\rho}{b}$$

In order to relate C_1 to the sources in our problem, we may take the curl of **A**,

$$\nabla\times\mathbf{A} = -\frac{\partial A_z}{\partial\rho}\mathbf{a}_\phi = -\frac{C_1}{\rho}\mathbf{a}_\phi = \mathbf{B}$$

obtain **H**,

$$\mathbf{H} = -\frac{C_1}{\mu_0\rho}\mathbf{a}_\phi$$

and evaluate the line integral,

$$\oint \mathbf{H}\cdot d\mathbf{L} = I = \int_0^{2\pi} -\frac{C_1}{\mu_0\rho}\mathbf{a}_\phi\cdot\rho\,d\phi\,\mathbf{a}_\phi = -\frac{2\pi C_1}{\mu_0}$$

Thus

$$C_1 = -\frac{\mu_0 I}{2\pi}$$

or

$$A_z = \frac{\mu_0 I}{2\pi}\ln\frac{b}{\rho} \tag{66}$$

and

$$H_\phi = \frac{I}{2\pi\rho}$$

as before. A plot of A_z versus ρ for $b = 5a$ is shown in Fig. 8.20; the decrease of $|\mathbf{A}|$ with distance from the concentrated current source which the inner conductor represents is evident. The results of Prob. D8.9 have also been added to the graph of Fig. 8.20. The extension of the curve into the outer conductor is left as Prob. 8.43.

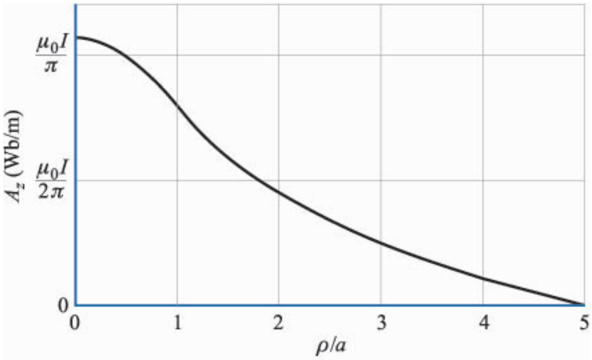


FIGURE 8.20
The vector magnetic potential is shown within the inner conductor and in the region between conductors for a coaxial cable with $b = 5a$ carrying I in the $\mathbf{a}_z$ direction, $A_z = 0$ is arbitrarily selected at $\rho = b$.

It is also possible to find A_z between conductors by applying a process some of us informally call "uncurling." That is, we know **H** or **B** for the coax, and we may therefore select the ϕ component of $\nabla \times \mathbf{A} = \mathbf{B}$ and integrate to obtain A_z. Try it, you'll like it!

✔ **D8.10.** Equation (66) is obviously also applicable to the exterior of any conductor of circular cross section carrying a current I in the $\mathbf{a}_z$ direction in free space. The zero reference is arbitrarily set at $\rho = b$. Now consider two conductors, each of 1-cm radius, parallel to the z axis with their axes lying in the $x = 0$ plane. One conductor whose axis is at $(0, 4\,\text{cm}, z)$ carries 12 A in the $\mathbf{a}_z$ direction; the other axis is at $(0, -4\,\text{cm}, z)$ and carries 12 A in the $-\mathbf{a}_z$ direction. Each current has its zero reference for **A** located 4 cm from its axis. Find the total **A** field at: $(0, 0, z)$; (*b*) $(0, 8\,\text{cm}, z)$; (*c*) $(4\,\text{cm}, 4\,\text{cm}, z)$; (*d*) $(2\,\text{cm}, 4\,\text{cm}, z)$.

Ans. 0; 2.20 μ Wb/m; 0; 8.93 μ Wb/m

SUGGESTED REFERENCES

1. Boast, W. B.: (see Suggested References for Chap. 2). The scalar magnetic potential is defined on p. 220, and its use in mapping magnetic fields is discussed on p. 444.
2. Jordan, E. C., and K. G. Balmain: "Electromagnetic Waves and Radiating Systems," 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968. Vector magnetic potential is discussed on pp. 90–96.
3. Paul, C. R., K. W. Whites, and S. Y. Nasar: "Introduction to Electromagnetic Fields," 3rd ed., McGraw-Hill Book Company, New York, 1998. The vector magnetic potential is presented on pp. 216–220.
4. Skilling, H. H.: (see Suggested References for Chap. 3). The "paddle wheel" is introduced on pp. 23–25.

PROBLEMS

8.1 (*a*) Find $\mathbf{H}$ in cartesian components at $P(2, 3, 4)$ if there is a current filament on the z axis carrying 8 mA in the $\mathbf{a}_z$ direction. (*b*) Repeat if the filament is located at $x = -1$, $y = 2$. (*c*) Find $\mathbf{H}$ if both filaments are present.

8.2 A current filament of $3\mathbf{a}_x$ A lies along the x axis. Find $\mathbf{H}$ in cartesian components at $P(-1, 3, 2)$.

8.3 Two semi-infinite filaments on the z axis lie in the regions $-\infty < z < a$ and $a < z < \infty$. Each carries a current I in the $\mathbf{a}_z$ direction. (*a*) Calculate $\mathbf{H}$ as a function of ρ and ϕ at $z = 0$. (*b*) What value of a will cause the magnitude of $\mathbf{H}$ at $\rho = 1$, $z = 0$, to be half the value obtained for an infinite filament?

8.4 (*a*) A filament is formed into a circle of radius a, centered at the origin in the plane $z = 0$. It carries a current I in the $\mathbf{a}_\phi$ direction. Find $\mathbf{H}$ at the origin. (*b*) A filament of the same length is shaped into a square in the $z = 0$ plane. The sides are parallel to the coordinate axes and a current I flows in the general $\mathbf{a}_\phi$ direction. Again find $\mathbf{H}$ at the origin.

8.5 The parallel filamentary conductors shown in Fig. 8.21 lie in free space. Plot $|\mathbf{H}|$ versus y, $-4 < y < 4$, along the line $x = 0$, $z = 2$.

8.6 (*a*) A current filament I is formed into a circle, $\rho = a$, in the $z = z'$ plane. Find H_z at $P(0, 0, z)$ if I flows in the $\mathbf{a}_\phi$ direction. (*b*) Find H_z at P caused by a uniform surface current density $\mathbf{K} = K_0\mathbf{a}_\phi$, flowing on the cylindrical surface, $\rho = a$, $0 < z < h$. The results of part (*a*) should help.

8.7 Given points $C(5, -2, 3)$ and $P(4, -1, 2)$, a current element $Id\mathbf{L} = 10^{-4}(4, -3, 1)$ A · m at C produces a field $d\mathbf{H}$ at P. (*a*) Specify the direction of $d\mathbf{H}$ by a unit vector $\mathbf{a}_H$. (*b*) Find $|d\mathbf{H}|$. (*c*) What direction $\mathbf{a}_1$ should $Id\mathbf{L}$ have at C so that $d\mathbf{H} = 0$?

8.8 For the finite-length current element on the z axis, as shown in Fig. 8.5, use the Biot-Savart law to derive Eq. (9) of Sec. 8.1.

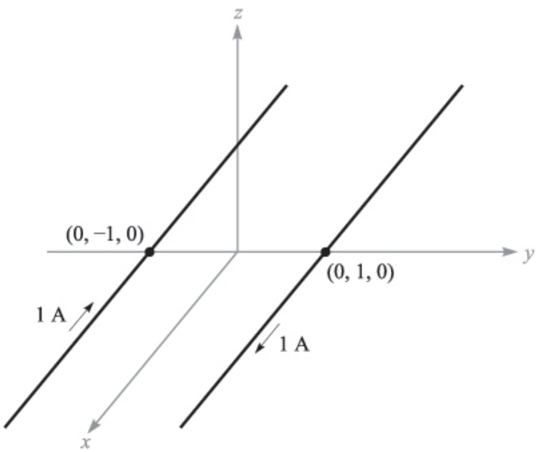


FIGURE 8.21
See Prob. 8.5.

8.9 A current sheet $\mathbf{K} = 8\mathbf{a}_x$ A/m flows in the region $-2 < y < 2$ m in the plane $z = 0$. Calculate $\mathbf{H}$ at $P(0, 0, 3)$.

8.10 Let a filamentary current of 5 mA be directed from infinity to the origin on the positive z axis and then back out to infinity on the positive x axis. Find $\mathbf{H}$ at $P(0, 1, 0)$.

8.11 An infinite filament on the z axis carries 20π mA in the $\mathbf{a}_z$ direction. Three uniform cylindrical current sheets are also present: 400 mA/m at $\rho = 1$ cm, -250 mA/m at $\rho = 2$ cm, and -300 mA/m at $\rho = 3$ cm. Calculate H_ϕ at $\rho = 0.5$, 1.5, 2.5, and 3.5 cm.

8.12 In Fig. 8.22, let the regions $0 < z < 0.3$ m and $0.7 < z < 1.0$ m be conducting slabs carrying uniform current densities of 10 A/m^2 in opposite directions as shown. Find $\mathbf{H}$ at $z =$: (*a*) -0.2; (*b*) 0.2 (*c*) 0.4; (*d*) 0.75; (*e*) 1.2 m.

8.13 A hollow cylindrical shell of radius a is centered on the z axis and carries a uniform surface current density of $K_a\mathbf{a}_\phi$. (*a*) Show that H is not a function of ϕ or z. (*b*) Show that H_ϕ and H_ρ are everywhere zero. (*c*) Show that $H_z = 0$ for $\rho > a$. (*d*) Show that $H_z = K_a$ for $\rho < a$. (*e*) A second shell, $\rho = b$, carries a current $K_b\mathbf{a}_\phi$. Find $\mathbf{H}$ everywhere.

8.14 A toroid having a cross section of rectangular shape is defined by the following surfaces: the cylinders $\rho = 2$ cm and $\rho = 3$ cm, and the planes $z = 1$ cm and $z = 2.5$ cm. The toroid carries a surface current density of $-50\mathbf{a}_z$ A/m on the surface $\rho = 3$ cm. Find $\mathbf{H}$ at the point $P(\rho, \phi, z)$: (*a*) $P_A(1.5\text{ cm}, 0, 2\text{ cm})$; (*b*) $P_B(2.1\text{ cm}, 0, 2\text{ cm})$; (*c*) $P_C(2.7\text{ cm}, \pi/2, 2\text{ cm})$; (*d*) $3.5\text{ cm}, \pi/2, 2\text{ cm})$.

8.15 Assume that there is a region with cylindrical symmetry in which the conductivity is given by $\sigma = 1.5e^{-150\rho}$ kS/m. An electric field of $30\mathbf{a}_z$ V/m is present. (*a*) Find $\mathbf{J}$. (*b*) Find the total current crossing the surface $\rho < \rho_0$, $z = 0$, all ϕ. (*c*) Make use of Ampere's circuital law to find $\mathbf{H}$.

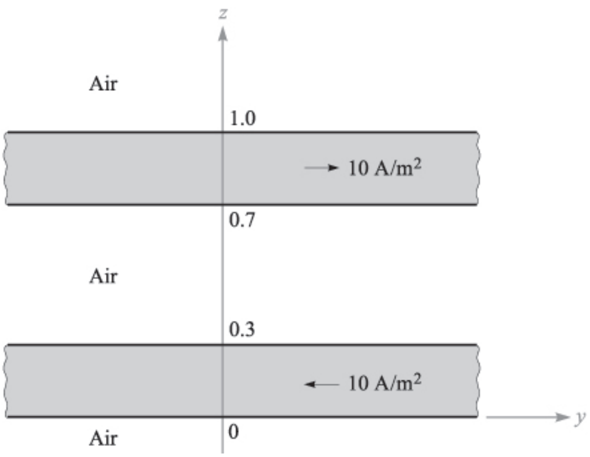


FIGURE 8.22
See Prob. 8.12.

8.16 The cylindrical shell, $2\,\text{mm} < \rho < 3\,\text{mm}$, carries a uniformly distributed total current of 8 A in the $-\mathbf{a}_z$ direction, and a filament on the z axis carries 8 A in the $\mathbf{a}_z$ direction. Find $\mathbf{H}$ everywhere.

8.17 A current filament on the z axis carries a current of 7 mA in the $\mathbf{a}_z$ direction, and current sheets of $0.5\mathbf{a}_z$ A/m and $-0.2\mathbf{a}_z$ A/m are located at $\rho = 1$ cm and $\rho = 0.5$ cm, respectively. Calculate $\mathbf{H}$ at $\rho =$: (*a*) 0.5 cm; (*b*) 1.5 cm; (*c*) 4 cm; (*d*) What current sheet should be located at $\rho = 4$ cm so that $\mathbf{H} = 0$ for all $\rho > 4$ cm?

8.18 Current density is distributed as follows: $\mathbf{J} = 0$ for $|y| > 2$ m, $\mathbf{J} = 8y\mathbf{a}_z$ A/m^2 for $|y| < 1$ m, $\mathbf{J} = 8(2-y)\mathbf{a}_z$ A/m^2 for $1 < y < 2$ m, $\mathbf{J} = -8(2+y)\mathbf{a}_z$ A/m^2 for $-2 < y < -1$ m. Use symmetry and Ampere's law to find $\mathbf{H}$ everywhere.

8.19 Calculate $\nabla \times [\nabla(\nabla \cdot \mathbf{G})]$ if $\mathbf{G} = 2x^2yz\mathbf{a}_x - 20y\mathbf{a}_y + (x^2 - z^2)\mathbf{a}_z$.

8.20 The magnetic field intensity is given in the square region $x = 0$, $0.5 < y < 1$, $1 < z < 1.5$ by $\mathbf{H} = z^2\mathbf{a}_x + x^3\mathbf{a}_y + y^4\mathbf{a}_z$ A/m. (*a*) Evaluate $\oint \mathbf{H} \cdot d\mathbf{L}$ about the perimeter of the square region. (*b*) Find $\nabla \times \mathbf{H}$. (*c*) Calculate $(\nabla \times \mathbf{H})_x$ at the center of the region. (*d*) Does $(\nabla \times \mathbf{H})_x = [\oint \mathbf{H} \cdot d\mathbf{L}]/\text{Area enclosed}$?

8.21 Points A, B, C, D, E, and F are each 2 mm from the origin on the coordinate axis indicated in Fig. 8.23. The value of $\mathbf{H}$ at each point is given. Calculate an approximate value for $\nabla \times \mathbf{H}$ at the origin.

8.22 In the cylindrical region $\rho \le 0.6$ mm, $H_\phi = \dfrac{2}{\rho} + \dfrac{\rho}{2}$ A/m, while $H_\phi = \dfrac{3}{\rho}$ A/m for $\rho > 0.6$ mm. (*a*) Determine $\mathbf{J}$ for $\rho < 0.6$ mm. (*b*) Determine $\mathbf{J}$ for $\rho > 0.6$ mm. (*c*) Is there a filamentary current at $\rho = 0$? If so, what is its value? (*d*) What is $\mathbf{J}$ at $\rho = 0$?

8.23 Given the field $\mathbf{H} = 20\rho^2\mathbf{a}_\phi$ A/m: (*a*) determine the current density $\mathbf{J}$; (*b*) integrate $\mathbf{J}$ over the circular surface $\rho = 1$, $0 < \phi < 2\pi$, $z = 0$, to determine the total current passing through that surface in the $\mathbf{a}_z$ direction; (*c*) find the total current once more, this time by a line integral around the circular path $\rho = 1$, $0 < \phi < 2\pi$, $z = 0$.

Point	H (A/m)		
A	$11.34\mathbf{a}_x$	$-13.78\mathbf{a}_y$	$+14.21\mathbf{a}_z$
B	$10.68\mathbf{a}_x$	$-12.19\mathbf{a}_y$	$+15.82\mathbf{a}_z$
C	$10.49\mathbf{a}_x$	$-12.19\mathbf{a}_y$	$+15.69\mathbf{a}_z$
D	$11.49\mathbf{a}_x$	$-13.78\mathbf{a}_y$	$+14.35\mathbf{a}_z$
E	$11.11\mathbf{a}_x$	$-13.88\mathbf{a}_y$	$+15.10\mathbf{a}_z$
F	$10.88\mathbf{a}_x$	$-13.10\mathbf{a}_y$	$+14.90\mathbf{a}_z$

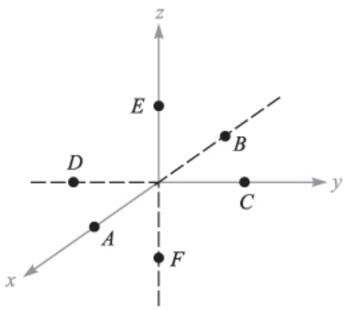


FIGURE 8.23
See Prob. 8.21.

8.24 Evaluate both sides of Stokes' theorem for the field $\mathbf{G} = 10\sin\theta\mathbf{a}_\phi$ and the surface $r = 3$, $0 \le \theta \le 90^\circ$, $0 \le \phi \le 90^\circ$. Let the surface have the $\mathbf{a}_r$ direction.

8.25 Given the field $\mathbf{H} = \frac{1}{2}\cos\frac{\phi}{2}\mathbf{a}_\rho - \sin\frac{\phi}{2}\mathbf{a}_\phi$ A/m, evaluate both sides of Stokes' theorem for the path formed by the intersection of the cylinder $\rho = 3$ and the plane $z = 2$, and for the surface defined by $\rho = 3$, $0 \le z \le 2$, and $z = 0$, $0 \le \rho \le 3$.

8.26 Let $\mathbf{G} = 15r\mathbf{a}_\phi$. (*a*) Determine $\oint \mathbf{G}\cdot d\mathbf{L}$ for the circular path $r = 5$, $\theta = 25^\circ$, $0 \le \phi \le 2\pi$. Evaluate $\int_S (\nabla\times\mathbf{G})\cdot d\mathbf{S}$ over the spherical cap $r = 5$, $0 \le \theta \le 25^\circ$, $0 \le \phi \le 2\pi$.

8.27 The magnetic field intensity is given in a certain region of space as $\mathbf{H} = \frac{x+2y}{z^2}\mathbf{a}_y + \frac{2}{z}\mathbf{a}_z$ A/m. (*a*) Find $\nabla\times\mathbf{H}$. (*b*) Find $\mathbf{J}$. (*c*) Use $\mathbf{J}$ to find the total current passing through the surface $z = 4$, $1 \le x \le 2$, $3 \le z \le 5$, in the $\mathbf{a}_z$ direction. (*d*) Show that the same result is obtained using the other side of Stokes' theorem.

8.28 Given $\mathbf{H} = (3r^2/\sin\theta)\mathbf{a}_\theta + 54r\cos\theta\mathbf{a}_\phi$ A/m in free space: (*a*) find the total current in the $\mathbf{A}_\theta$ direction through the conical surface $\theta = 20^\circ$, $0 \le \phi \le 2\pi$, $0 \le r \le 5$, by whichever side of Stokes' theorem you like the best. (*b*) Check the result by using the other side of Stokes' theorem.

8.29 A long straight nonmagnetic conductor of 0.2-mm radius carries a uniformly distributed current of 2 A dc. (*a*) Find $\mathbf{J}$ within the conductor. (*b*) Use Ampère's circuital law to find $\mathbf{H}$ and $\mathbf{B}$ within the conductor. (*c*) Show that $\nabla\times\mathbf{H} = \mathbf{J}$ within the conductor. (*d*) Find $\mathbf{H}$ and $\mathbf{B}$ within the conductor. (*e*) Show that $\nabla\times\mathbf{H} = \mathbf{J}$ outside the conductor.

8.30 A solid nonmagnetic conductor of circular cross section has a radius of 2 mm. The conductor is inhomogeneous, with $\sigma = 10^6(1 + 10^6\rho^2)$ S/m. If the conductor is 1 m in length and has a voltage of 1 mV between its ends, find: (*a*) H; (*b*) the total magnetic flux inside the conductor.

8.31 The cylindrical shell defined by $1\text{ cm} < \rho < 1.4\text{ cm}$ consists of a nonmagnetic conducting material and carries a total current of 50 A in the $\mathbf{a}_z$ direction. Find the total magnetic flux crossing the plane $\phi = 0$, $0 < z < 1$: (*a*) $0 < \rho < 1.2$ cm; (*b*) $1.4\text{ cm} < \rho < 1.4$ cm; (*c*) $1.4\text{ cm} < \rho < 20$ cm.

8.32 The free-space region defined by $1 < z < 4$ cm and $2 < \rho < 3$ cm is a toroid of rectangular cross section. Let the surface at $\rho = 3$ cm carry a surface current $\mathbf{K} = 2\mathbf{a}_z$ kA/m. (*a*) Specify the currents on the surfaces at $\rho = 2$ cm, $z = 1$ cm, and $z = 4$ cm. (*b*) Find $\mathbf{H}$ everywhere. (*c*) Calculate the total flux within the toroid.

8.33 Use an expansion in cartesian coordinates to show that the curl of the gradient of any scalar field G is identically equal to zero.

8.34 A filamentary conductor on the z axis carries a current of 16 A in the $\mathbf{a}_z$ direction, a conducting shell at $\rho = 6$ carries a total current of 12 A in the $-\mathbf{a}_z$ direction, and another shell at $\rho = 10$ carries a total current of 4 A in

the $-\mathbf{a}_z$ direction. (*a*) Find $\mathbf{H}$ for $0 < \rho < 12$. (*b*) Plot H_ϕ versus ρ. (*c*) Find the total flux Φ crossing the surface $1 < \rho < 7$, $0 < z < 1$.

8.35 A current sheet, $\mathbf{K} = 20\mathbf{a}_z$ A/m, is located at $\rho = 2$, and a second sheet, $\mathbf{K} = -10\mathbf{a}_z$ A/m, is located at $\rho = 4$. (*a*) Let $V_m = 0$ at $P(\rho = 3, \phi = 0, z = 5)$ and place a barrier at $\phi = \pi$. Find $V_m(\rho, \phi, z)$ for $-\pi < \phi < \pi$. (*b*) Let $\mathbf{A} = 0$ at P and find $\mathbf{A}(\rho, \phi, z)$ for $2 < \rho < 4$.

8.36 Let $\mathbf{A} = (3y - z)\mathbf{a}_x + 2xz\mathbf{a}_y$ Wb/m in a certain region of free space. (*a*) Show that $\nabla \cdot \mathbf{A} = 0$. (*b*) At $P(2, -1, 3)$, find $\mathbf{A}$, $\mathbf{B}$, $\mathbf{H}$, and $\mathbf{J}$.

8.37 Let $N = 1000$, $I = 0.8$ A, $\rho_0 = 2$ cm, and $a = 0.8$ cm for the toroid shown in Fig. 8.12*b*. Find V_m in the interior of the toroid if $V_m = 0$ at $\rho = 2.5$ cm, $\phi = 0.3\pi$. Keep ϕ within the range $0 < \phi < 2\pi$.

8.38 The solenoid shown in Fig. 8.11*b* contains 400 turns, carries a current $I = 5$ A, has a length of 8 cm, and a radius $a = 1.2$ cm. (*a*) Find $\mathbf{H}$ within the solenoid. (*b*) If $V_m = 0$ at the origin, specify $V_m(\rho, \phi, z)$ inside the solenoid. (*c*) Let $\mathbf{A} = 0$ at the origin, and specify $\mathbf{A}(\rho, \phi, z)$ inside the solenoid if the medium is free space.

8.39 Planar current sheets of $\mathbf{K} = 30\mathbf{a}_z$ A/m and $-30\mathbf{a}_z$ A/m are located in free space at $x = 0.2$ and $x = -0.2$, respectively. For the region $-0.2 < x < 0.2$: (*a*) find $\mathbf{H}$; (*b*) obtain an expression for V_m if $V_m = 0$ at $P(0.1, 0.2, 0.3)$; (*c*) find $\mathbf{B}$; (*d*) obtain an expression for $\mathbf{A}$ if $\mathbf{A} = 0$ at P.

8.40 Let $\mathbf{A} = (3y^2 - 2z)\mathbf{a}_x - 2x^2z\mathbf{a}_y + (x + 2y)\mathbf{a}_z$ Wb/m in free space. Find $\nabla \times \nabla \times \mathbf{A}$ at $P(-2, 3, -1)$.

8.41 Assume that $\mathbf{A} = 50\rho^2\mathbf{a}_z$ Wb/m in a certain region of free space. (*a*) Find $\mathbf{H}$ and $\mathbf{B}$. (*b*) Find $\mathbf{J}$. (*c*) Use $\mathbf{J}$ to find the total current crossing the surface $0 \le \rho \le 1$, $0 \le \phi < 2\pi$, $z = 0$. (*d*) Use the value of H_ϕ at $\rho = 1$ to calculate $\oint \mathbf{H} \cdot d\mathbf{L}$ for $\rho = 1$, $z = 0$.

8.42 Show that $\nabla_2(1/R_{12}) = -\nabla_1(1/R_{12}) = \mathbf{R}_{21}/R_{12}^3$.

8.43 Compute the vector magnetic potential within the outer conductor for the coaxial line whose vector magnetic potential is shown in Fig. 8.20 if the outer radius of the outer conductor is $7a$. Select the proper zero reference and sketch the results on the figure.

8.44 By expanding Eq. (58), Sec. 8.7, in cartesian coordinates, show that (59) is correct.

CHAPTER

9

MAGNETIC FORCES, MATERIALS, AND INDUCTANCE

The magnetic field quantities **H**, **B**, Φ, V_m, and **A** introduced in the last chapter were not given much physical significance. Each of these quantities is merely defined in terms of the distribution of current sources throughout space. If the current distribution is known, we should feel that **H**, **B**, and **A** are determined at every point in space, even though we may not be able to evaluate the defining integrals because of mathematical complexity.

We are now ready to undertake the second half of the magnetic field problem, that of determining the forces and torques exerted by the magnetic field on other charges. The electric field causes a force to be exerted on a charge which may be either stationary or in motion; we shall see that the steady magnetic field is capable of exerting a force only on a *moving* charge. This result appears reasonable; a magnetic field may be produced by moving charges and may exert forces on moving charges; a magnetic field cannot arise from stationary charges and cannot exert any force on a stationary charge.

This chapter initially considers the forces and torques on current-carrying conductors which may either be of a filamentary nature or possess a finite cross section with a known current density distribution. The problems associated with the motion of particles in a vacuum are largely avoided.

With an understanding of the fundamental effects produced by the magnetic field, we may then consider the varied types of magnetic materials, the analysis of elementary magnetic circuits, the forces on magnetic materials, and finally, the important electrical circuit concepts of self-inductance and mutual inductance.

9.1 FORCE ON A MOVING CHARGE

In an electric field the definition of the electric field intensity shows us that the force on a charged particle is

$$\boxed{\mathbf{F} = Q\mathbf{E}} \tag{1}$$

The force is in the same direction as the electric field intensity (for a positive charge) and is directly proportional to both $\mathbf{E}$ and Q. If the charge is in motion, the force at any point in its trajectory is then given by (1).

A charged particle in motion in a magnetic field of flux density $\mathbf{B}$ is found experimentally to experience a force whose magnitude is proportional to the product of the magnitudes of the charge Q, its velocity $\mathbf{v}$, and the flux density $\mathbf{B}$, and to the sine of the angle between the vectors $\mathbf{v}$ and $\mathbf{B}$. The direction of the force is perpendicular to both $\mathbf{v}$ and $\mathbf{B}$ and is given by a unit vector in the direction of $\mathbf{v} \times \mathbf{B}$. The force may therefore be expressed as

$$\boxed{\mathbf{F} = Q\mathbf{v} \times \mathbf{B}} \tag{2}$$

A fundamental difference in the effect of the electric and magnetic fields on charged particles is now apparent, for a force which is always applied in a direction at right angles to the direction in which the particle is proceeding can never change the magnitude of the particle velocity. In other words, the *acceleration* vector is always normal to the velocity vector. The kinetic energy of the particle remains unchanged, and it follows that the steady magnetic field is incapable of transferring energy to the moving charge. The electric field, on the other hand, exerts a force on the particle which is independent of the direction in which the particle is progressing and therefore effects an energy transfer between field and particle in general.

The first two problems at the end of this chapter illustrate the different effects of electric and magnetic fields on the kinetic energy of a charged particle moving in free space.

The force on a moving particle due to combined electric and magnetic fields is obtained easily by superposition,

$$\boxed{\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B})} \tag{3}$$

This equation is known as the *Lorentz force equation*, and its solution is required in determining electron orbits in the magnetron, proton paths in the cyclotron, plasma characteristics in a magnetohydrodynamic (MHD) generator, or, in general, charged-particle motion in combined electric and magnetic fields.

✔ **D9.1.** The point charge $Q = 18\,\text{nC}$ has a velocity of $5 \times 10^6\,\text{m/s}$ in the direction $\mathbf{a}_v = 0.04\mathbf{a}_x - 0.05\mathbf{a}_y + 0.2\mathbf{a}_z$. Calculate the magnitude of the force exerted on the charge by the field: (*a*) $\mathbf{B} = -3\mathbf{a}_x + 4\mathbf{a}_y + 6\mathbf{a}_z$ mT; (*b*) $\mathbf{E} = -3\mathbf{a}_x + 4\mathbf{a}_y + 6\mathbf{a}_z$ kV/m; (*c*) **B** and **E** acting together.

Ans. 124.6 μN; 140.6 μN; 187.8 μN

9.2 FORCE ON A DIFFERENTIAL CURRENT ELEMENT

The force on a charged particle moving through a steady magnetic field may be written as the differential force exerted on a differential element of charge,

$$d\mathbf{F} = dQ\,\mathbf{v} \times \mathbf{B} \tag{4}$$

Physically, the differential element of charge consists of a large number of very small discrete charges occupying a volume which, although small, is much larger than the average separation between the charges. The differential force expressed by (4) is thus merely the sum of the forces on the individual charges. This sum, or resultant force, is not a force applied to a single object. In an analogous way, we might consider the differential gravitational force experienced by a small volume taken in a shower of falling sand. The small volume contains a large number of sand grains, and the differential force is the sum of the forces on the individual grains within the small volume.

If our charges are electrons in motion in a conductor, however, we can show that the force is transferred to the conductor and that the sum of this extremely large number of extremely small forces is of practical importance. Within the conductor, electrons are in motion throughout a region of immobile positive ions which form a crystalline array giving the conductor its solid properties. A magnetic field which exerts forces on the electrons tends to cause them to shift position slightly and produces a small displacement between the centers of "gravity" of the positive and negative charges. The Coulomb forces between electrons and positive ions, however, tend to resist such a displacement. Any attempt to move the electrons, therefore, results in an attractive force between electrons and the positive ions of the crystalline lattice. The magnetic force is thus transferred to the crystalline lattice, or to the conductor itself. The Coulomb forces are so much greater than the magnetic forces in good conductors that the actual displacement of the electrons is almost immeasurable. The charge separation that does result, however, is disclosed by the presence of a slight potential difference across the conductor sample in a direction perpendicular to both the

magnetic field and the velocity of the charges. The voltage is known as the *Hall voltage*, and the effect itself is called the *Hall effect*.

Fig. 9.1 illustrates the direction of the Hall voltage for both positive and negative charges in motion. In Fig. 9.1*a*, $\mathbf{v}$ is in the $-\mathbf{a}_x$ direction, $\mathbf{v} \times \mathbf{B}$ is in the $\mathbf{a}_y$ direction, and Q is positive, causing $\mathbf{F}_Q$ to be in the $\mathbf{a}_y$ direction; thus, the positive charges move to the right. In Figure 9.1*b*, $\mathbf{v}$ is now in the $+\mathbf{a}_x$ direction, $\mathbf{B}$ is still in the $\mathbf{a}_z$ direction, $\mathbf{v} \times \mathbf{B}$ is in the $-\mathbf{a}_y$ direction, and Q is negative; thus $\mathbf{F}_Q$ is again in the $\mathbf{a}_y$ direction. Hence, the negative charges end up at the right edge. Equal currents provided by holes and electrons in semiconductors can therefore be differentiated by their Hall voltages. This is one method of determining whether a given semiconductor is *n*-type or *p*-type.

Devices employ the Hall effect to measure the magnetic flux density and, in some applications where the current through the device can be made proportional to the magnetic field across it, to serve as electronic wattmeters, squaring elements, and so forth.

Returning to (4), we may therefore say that if we are considering an element of moving charge in an electron beam, the force is merely the sum of the forces on the individual electrons in that small volume element, but if we are considering an element of moving charge within a conductor, the total force is applied to the solid conductor itself. We shall now limit our attention to the forces on current-carrying conductors.

In Chap. 5 we defined convection current density in terms of the velocity of the volume charge density,

$$\boxed{\mathbf{J} = \rho_v \mathbf{v}}$$

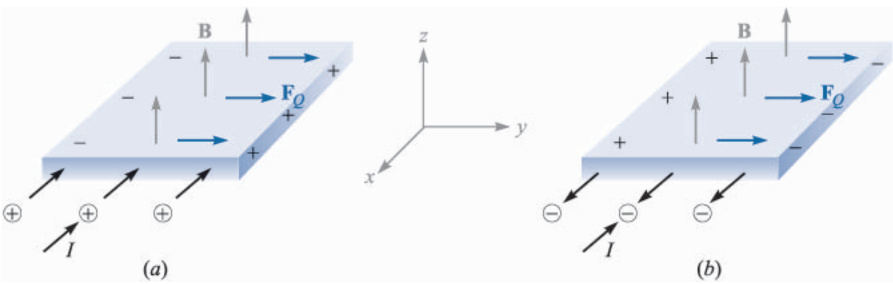


FIGURE 9.1
Equal currents directed into the material are provided by positive charges moving inward in (*a*) and negative charges moving outward in (*b*). The two cases can be distinguished by oppositely directed Hall voltages, as shown.

The differential element of charge in (4) may also be expressed in terms of volume charge density,[1]

$$dQ = \rho_v dv$$

Thus

$$d\mathbf{F} = \rho_v dv\, \mathbf{v} \times \mathbf{B}$$

or

$$\boxed{d\mathbf{F} = \mathbf{J} \times \mathbf{B}\, dv} \tag{5}$$

We saw in the previous chapter that $\mathbf{J}\, dv$ may be interpreted as a differential current element; that is,

$$\mathbf{J}\, dv = \mathbf{K}\, dS = I\, d\mathbf{L}$$

and thus the Lorentz force equation may be applied to surface current density,

$$\boxed{d\mathbf{F} = \mathbf{K} \times \mathbf{B}\, dS} \tag{6}$$

or to a differential current filament,

$$\boxed{d\mathbf{F} = I\, d\mathbf{L} \times \mathbf{B}} \tag{7}$$

Integrating (5), (6), or (7) over a volume, a surface which may be either open or closed (why?), or a closed path, respectively, leads to the integral formulations

$$\mathbf{F} = \int_{\text{vol}} \mathbf{J} \times \mathbf{B}\, dv \tag{8}$$

$$\mathbf{F} = \int_S \mathbf{K} \times \mathbf{B}\, dS \tag{9}$$

and

$$\boxed{\mathbf{F} = \oint I\, d\mathbf{L} \times \mathbf{B} = -I \oint \mathbf{B} \times d\mathbf{L}} \tag{10}$$

One simple result is obtained by applying (7) or (10) to a straight conductor in a uniform magnetic field,

[1] Remember that dv is a differential volume element and not a differential increase in velocity.

$$\boxed{\mathbf{F} = I\mathbf{L} \times \mathbf{B}} \tag{11}$$

The magnitude of the force is given by the familiar equation

$$F = BIL \sin\theta \tag{12}$$

where θ is the angle between the vectors representing the direction of the current flow and the direction of the magnetic flux density. Equation (11) or (12) applies only to a portion of the closed circuit, and the remainder of the circuit must be considered in any practical problem.

Example 9.1

As a numerical example of these equations, consider Fig. 9.2. We have a square loop of wire in the $z = 0$ plane carrying 2 mA in the field of an infinite filament on the y axis, as shown. We desire the total force on the loop.

Solution. The field produced in the plane of the loop by the straight filament is

$$\mathbf{H} = \frac{I}{2\pi x}\mathbf{a}_z = \frac{15}{2\pi x}\mathbf{a}_z \quad \text{A/m}$$

Therefore,

$$\mathbf{B} = \mu_0 \mathbf{H} = 4\pi \times 10^{-7}\mathbf{H} = \frac{3 \times 10^{-6}}{x}\mathbf{a}_z \quad \text{T}$$

We use the integral form (10),

$$\mathbf{F} = -I \oint \mathbf{B} \times d\mathbf{L}$$

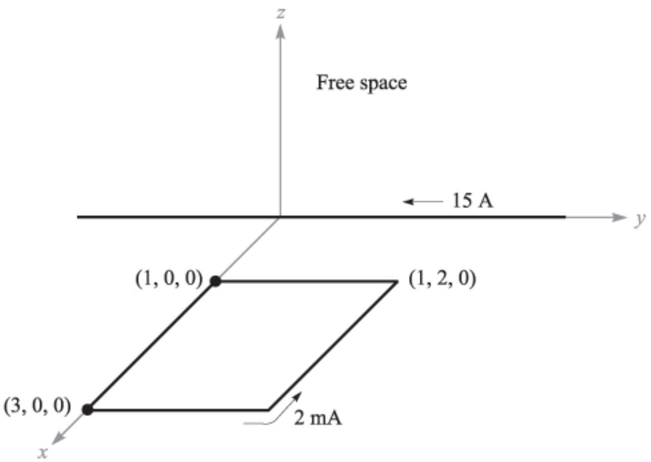


FIGURE 9.2
A square loop of wire in the xy plane carrying 2 mA is subjected to a nonuniform **B** field.

Let us assume a rigid loop so that the total force is the sum of the forces on the four sides. Beginning with the left side:

$$\mathbf{F} = -2 \times 10^{-3} \times 3 \times 10^{-6} \left[\int_{x=1}^{3} \frac{\mathbf{a}_z}{x} \times dx\,\mathbf{a}_x + \int_{y=0}^{2} \frac{\mathbf{a}_z}{3} \times dy\,\mathbf{a}_y \right.$$

$$\left. + \int_{x=3}^{1} \frac{\mathbf{a}_z}{x} \times dx\,\mathbf{a}_x + \int_{y=2}^{0} \frac{\mathbf{a}_z}{1} \times dy\,\mathbf{a}_y \right]$$

$$= -6 \times 10^{-9} \left[\ln x \Big|_1^3 \mathbf{a}_y + \frac{1}{3} y \Big|_0^2 (-\mathbf{a}_x) + \ln x \Big|_3^1 \mathbf{a}_y + y \Big|_2^0 (-\mathbf{a}_x) \right]$$

$$= -6 \times 10^{-9} \left[(\ln 3)\mathbf{a}_y - \frac{2}{3}\mathbf{a}_x + \left(\ln \frac{1}{3} \right) \mathbf{a}_y + 2\mathbf{a}_x \right]$$

$$= -8\mathbf{a}_x \quad \text{pN}$$

Thus, the net force on the loop is in the $-\mathbf{a}_x$ direction.

✔ **D9.2.** The field $\mathbf{B} = -2\mathbf{a}_x + 3\mathbf{a}_y + 4\mathbf{a}_z$ mT is present in free space. Find the vector force exerted on a straight wire carrying 12 A in the $\mathbf{a}_{AB}$ direction, given $A(1, 1, 1)$ and: (*a*) $B(2, 1, 1)$; (*b*) $B(3, 5, 6)$.

Ans. $-48\mathbf{a}_y + 36\mathbf{a}_z$ mN; $12\mathbf{a}_x - 216\mathbf{a}_y + 168\mathbf{a}_z$ mN

✔ **D9.3.** The semiconductor sample shown in Fig. 9.1 is *n*-type silicon, having a rectangular cross section of 0.9 mm by 1.1 cm, and a length of 1.3 cm. Assume the electron and hole mobilities are 0.13 and 0.03 $\mathrm{m^2/V{\cdot}s}$, respectively, at the operating temperature. Let $B = 0.07$ T and the electric field intensity in the direction of the current flow be 800 V/m. Find the magnitude of: (*a*) the voltage across the sample length; (*b*) the drift velocity; (*c*) the transverse force per coulomb of moving charge caused by B; (*d*) the transverse electric field intensity; (*e*) the Hall voltage.

Ans. 10.40 V; 104.0 m/s; 7.28 N/C; 7.28 V/m; 80.1 mV

9.3 FORCE BETWEEN DIFFERENTIAL CURRENT ELEMENTS

The concept of the magnetic field was introduced to break into two parts the problem of finding the interaction of one current distribution on a second current distribution. It is possible to express the force on one current element directly in terms of a second current element without finding the magnetic field. Since we claimed that the magnetic-field concept simplifies our work, it then behooves us to show that avoidance of this intermediate step leads to more complicated expressions.

The magnetic field at point 2 due to a current element at point 1 was found to be

$$d\mathbf{H}_2 = \frac{I_1 d\mathbf{L}_1 \times \mathbf{a}_{R12}}{4\pi R_{12}^2}$$

Now, the differential force on a differential current element is

$$d\mathbf{F} = I\,d\mathbf{L} \times \mathbf{B}$$

and we apply this to our problem by letting $\mathbf{B}$ be $d\mathbf{B}_2$ (the differential flux density at point 2 caused by current element 1), by identifying $I\,d\mathbf{L}$ as $I_2 d\mathbf{L}_2$, and by symbolizing the differential amount of our differential force on element 2 as $d(d\mathbf{F}_2)$:

$$d(d\mathbf{F}_2) = I_2 d\mathbf{L}_2 \times d\mathbf{B}_2$$

Since $d\mathbf{B}_2 = \mu_0 d\mathbf{H}_2$, we obtain the force between two differential current elements,

$$d(d\mathbf{F}_2) = \mu_0 \frac{I_1 I_2}{4\pi R_{12}^2} d\mathbf{L}_2 \times (d\mathbf{L}_1 \times \mathbf{a}_{R12}) \tag{13}$$

Example 9.2

As an example that illustrates the use (and misuse) of these results, consider the two differential current elements shown in Fig. 9.3. We seek the differential force on $d\mathbf{L}_2$.

Solution. We have $I_1 d\mathbf{L}_1 = -3\mathbf{a}_y$ A·m at $P_1(5, 2, 1)$, and $I_2 d\mathbf{L}_2 = -4\mathbf{a}_z$ A·m at $P_2(1, 8, 5)$. Thus, $\mathbf{R}_{12} = -4\mathbf{a}_x + 6\mathbf{a}_y + 4\mathbf{a}_z$, and we may substitute these data into (13),

$$d(d\mathbf{F}_2) = \frac{4\pi 10^{-7}}{4\pi} \frac{(-4\mathbf{a}_x) \times [(-3\mathbf{a}_y) \times (-4\mathbf{a}_x + 6\mathbf{a}_y + 4\mathbf{a}_z)]}{(16 + 36 + 16)^{1.5}}$$
$$= 8.56\mathbf{a}_y \quad \text{nN}$$

Many chapters ago when we discussed the force exerted by one point charge on another point charge, we found that the force on the first charge was the negative of that on the second. That is, the total force on the system

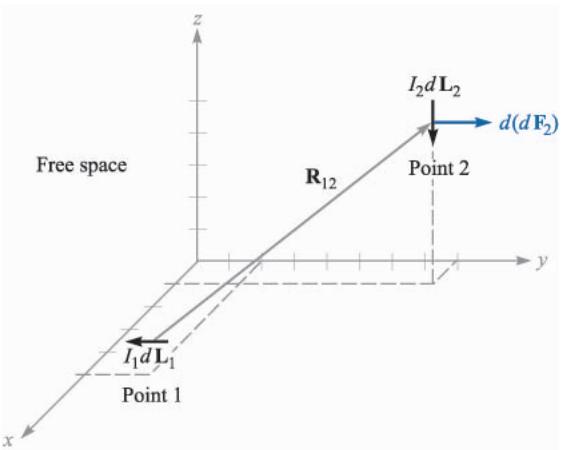


FIGURE 9.3
Given $P_1(5, 2, 1)$, $P_2(1, 8, 5)$, $I_1\,d\mathbf{L}_1 = -3\mathbf{a}_y$ A·m, and $I_2\,d\mathbf{L}_2 = -4\mathbf{a}_z$ A·m, the force on $I_2\,d\mathbf{L}_2$ is 8.56 nN in the $\mathbf{a}_y$ direction.

was zero. This is not the case with the differential current elements, and $d(d\mathbf{F}_1) = -12.84\mathbf{a}_z$ nN in the example above. The reason for this different behavior lies with the nonphysical nature of the current element. Whereas point charges may be approximated quite well by small charges, the continuity of current demands that a complete circuit be considered. This we shall now do.

The total force between two filamentary circuits is obtained by integrating twice:

$$\mathbf{F}_2 = \mu_0 \frac{I_1 I_2}{4\pi} \oint \left[d\mathbf{L}_2 \times \oint \frac{d\mathbf{L}_1 \times \mathbf{a}_{R12}}{R_{12}^2} \right] \\ = \mu_0 \frac{I_1 I_2}{4\pi} \oint \left[\oint \frac{\mathbf{a}_{R12} \times d\mathbf{L}_1}{R_{12}^2} \right] \times d\mathbf{L}_2 \tag{14}$$

Equation (14) is quite formidable, but the familiarity gained in the last chapter with the magnetic field should enable us to recognize the inner integral as the integral necessary to find the magnetic field at point 2 due to the current element at point 1.

Although we shall only give the result, it is not very difficult to make use of (14) to find the force of repulsion between two infinitely long, straight, parallel, filamentary conductors with separation d, and carrying equal but opposite currents I, as shown in Fig. 9.4. The integrations are simple, and most errors are made in determining suitable expressions for $\mathbf{a}_{R12}$, $d\mathbf{L}_1$, and $d\mathbf{L}_2$. However, since the magnetic field intensity at either wire caused by the other is already known to be $I/(2\pi d)$, it is readily apparent that the answer is a force of $\mu_0 I^2/(2\pi d)$ newtons per meter length.

✔ **D9.4.** Two differential current elements, $I_1 \Delta\mathbf{L}_1 = 3 \times 10^{-6}\mathbf{a}_y$ A · m at $P_1(1, 0, 0)$ and $I_2 \Delta\mathbf{L}_2 = 3 \times 10^{-6}(-0.5\mathbf{a}_x + 0.4\mathbf{a}_y + 0.3\mathbf{a}_z)$ A · m at $P_2(2, 2, 2)$, are located in free space. Find the vector force exerted on: (*a*) $I_2 \Delta\mathbf{L}_2$ by $I_1 \Delta\mathbf{L}_1$; (*b*) $I_1 \Delta\mathbf{L}_1$ by $I_2 \Delta\mathbf{L}_2$.

Ans. $(-1.333\mathbf{a}_x + 0.333\mathbf{a}_y - 2.67\mathbf{a}_z)10^{-20}$ N; $(4.33\mathbf{a}_x + 0.667\mathbf{a}_z)10^{-20}$ N

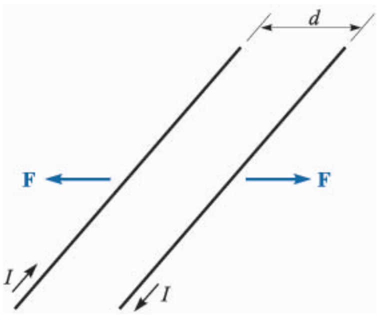


FIGURE 9.4
Two infinite parallel filaments with separation d and equal but opposite currents I experience a repulsive force of $\mu_0 I^2/(2\pi d)$ N/m.

9.4 FORCE AND TORQUE ON A CLOSED CIRCUIT

We have already obtained general expressions for the forces exerted on current systems. One special case is easily disposed of, for if we take our relationship for the force on a filamentary closed circuit, as given by Eq. (10), Sec. 9.2,

$$\mathbf{F} = -I \oint \mathbf{B} \times d\mathbf{L}$$

and assume a *uniform* magnetic flux density, then **B** may be removed from the integral:

$$\mathbf{F} = -I\mathbf{B} \times \oint d\mathbf{L}$$

However, we discovered during our investigation of closed line integrals in an electrostatic potential field that $\oint d\mathbf{L} = 0$, and therefore the force on a closed filamentary circuit in a uniform magnetic field is zero.

If the field is not uniform, the total force need not be zero.

This result for uniform fields does not have to be restricted to filamentary circuits only. The circuit may contain surface currents or volume current density as well. If the total current is divided into filaments, the force on each one is zero, as we showed above, and the total force is again zero. Therefore any real closed circuit carrying direct currents experiences a total vector force of zero in a uniform magnetic field.

Although the force is zero, the torque is generally not equal to zero.

In defining the *torque*, or *moment*, of a force, it is necessary to consider both an origin at or about which the torque is to be calculated, as well as the point at which the force is applied. In Fig. 9.5*a*, we apply a force **F** at point *P*, and we

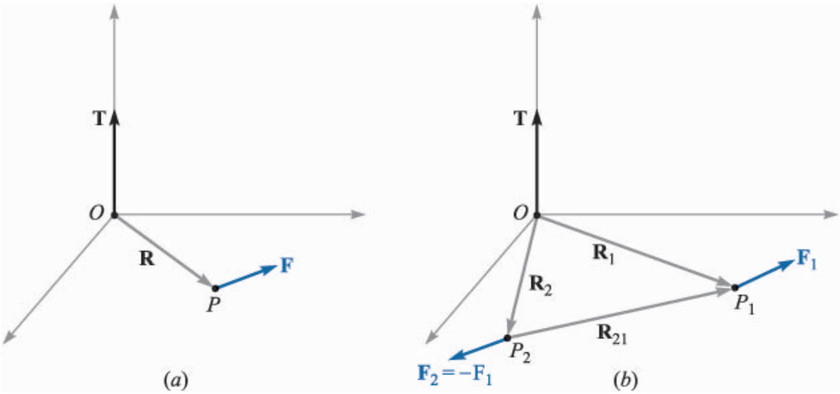


FIGURE 9.5
(*a*) Given a lever arm **R** extending from an origin *O* to a point *P* where force **F** is applied, the torque about *O* is $\mathbf{T} = \mathbf{R} \times \mathbf{F}$. (*b*) If $\mathbf{F}_2 = -\mathbf{F}_1$, then the torque $\mathbf{T} = \mathbf{R}_{21} \times \mathbf{F}_1$ is independent of the choice of origin for $\mathbf{R}_1$ and $\mathbf{R}_2$.

establish an origin at O with a rigid lever arm $\mathbf{R}$ extending from O to P. The torque about point O is a vector whose magnitude is the product of the magnitudes of $\mathbf{R}$, of $\mathbf{F}$, and of the sine of the angle between these two vectors. The direction of the vector torque $\mathbf{T}$ is normal to both the force $\mathbf{F}$ and lever arm $\mathbf{R}$ and is in the direction of progress of a right-handed screw as the lever arm is rotated into the force vector through the smaller angle. The torque is expressible as a cross product,

$$\mathbf{T} = \mathbf{R} \times \mathbf{F}$$

Now let us assume that two forces, $\mathbf{F}_1$ at P_1 and $\mathbf{F}_2$ at P_2, having lever arms $\mathbf{R}_1$ and $\mathbf{R}_2$ extending from a common origin O, as shown in Fig. 9.5*b*, are applied to an object of fixed shape and that the object does not undergo any translation. Then the torque about the origin is

$$\mathbf{T} = \mathbf{R}_1 \times \mathbf{F}_1 + \mathbf{R}_2 \times \mathbf{F}_2$$

where

$$\mathbf{F}_1 + \mathbf{F}_2 = 0$$

and therefore

$$\mathbf{T} = (\mathbf{R}_1 - \mathbf{R}_2) \times \mathbf{F}_1 = \mathbf{R}_{21} \times \mathbf{F}_1$$

The vector $\mathbf{R}_{21} = \mathbf{R}_1 - \mathbf{R}_2$ joins the point of application of $\mathbf{F}_2$ to that of $\mathbf{F}_1$ and is independent of the choice of origin for the two vectors $\mathbf{R}_1$ and $\mathbf{R}_2$. Therefore, the torque is also independent of the choice of origin, provided that the total force is zero. This may be extended to any number of forces.

Consider the application of a vertically upward force at the end of a horizontal crank handle on an elderly automobile. This cannot be the only applied force, for if it were, the entire handle would be accelerated in an upward direction. A second force, equal in magnitude to that exerted at the end of the handle, is applied in a downward direction by the bearing surface at the axis of rotation. For a 40-N force on a crank handle 0.3 m in length, the torque is 12 N·m. This figure is obtained regardless of whether the origin is considered to be on the axis of rotation (leading to 12 N·m plus 0 N·m), at the midpoint of the handle (leading to 6 N·m plus 6 N·m), or at some point not even on the handle or an extension of the handle.

We may therefore choose the most convenient origin, and this is usually on the axis of rotation and in the plane containing the applied forces if the several forces are coplanar.

With this introduction to the concept of torque, let us now consider the torque on a differential current loop in a magnetic field $\mathbf{B}$. The loop lies in the xy plane (Fig. 9.6); the sides of the loop are parallel to the x and y axes and are of length dx and dy. The value of the magnetic field at the center of the loop is taken as $\mathbf{B}_0$. Since the loop is of differential size, the value of $\mathbf{B}$ at all points on the loop may be taken as $\mathbf{B}_0$. (Why was this not possible in the discussion of curl and

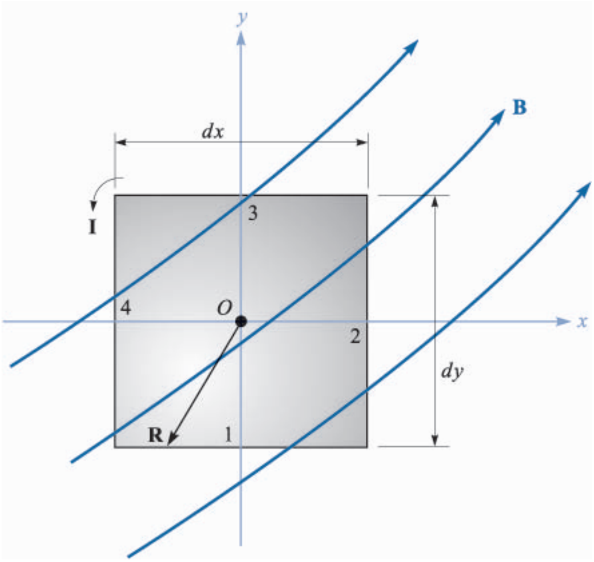


FIGURE 9.6
A differential current loop in a magnetic field **B**. The torque on the loop is $d\mathbf{T} = I(dx\,dy\mathbf{a}_z) \times \mathbf{B}_0 = I\,d\mathbf{S} \times \mathbf{B}$.

divergence?) The total force on the loop is therefore zero, and we are free to choose the origin for the torque at the center of the loop.

The vector force on side 1 is

$$d\mathbf{F}_1 = I\,dx\,\mathbf{a}_x \times \mathbf{B}_0$$

or

$$d\mathbf{F}_1 = I\,dx(B_{0y}\mathbf{a}_z - B_{0z}\mathbf{a}_y)$$

For this side of the loop the lever arm **R** extends from the origin to the midpoint of the side, $\mathbf{R}_1 = -\frac{1}{2}dy\,\mathbf{a}_y$, and the contribution to the total torque is

$$\begin{aligned} d\mathbf{T}_1 &= \mathbf{R}_1 \times d\mathbf{F}_1 \\ &= -\tfrac{1}{2}dy\,\mathbf{a}_y \times I\,dx(B_{0y}\mathbf{a}_z - B_{0z}\mathbf{a}_y) \\ &= -\tfrac{1}{2}dx\,dy\,IB_{0y}\mathbf{a}_x \end{aligned}$$

The torque contribution on side 3 is found to be the same,

$$\begin{aligned} d\mathbf{T}_3 &= \mathbf{R}_3 \times d\mathbf{F}_3 = \tfrac{1}{2}dy\,\mathbf{a}_y \times (-I\,dx\,\mathbf{a}_x \times \mathbf{B}_0) \\ &= -\tfrac{1}{2}dx\,dy\,IB_{0y}\mathbf{a}_x = d\mathbf{T}_1 \end{aligned}$$

and

$$d\mathbf{T}_1 + d\mathbf{T}_3 = -dx\,dy\,IB_{0y}\mathbf{a}_x$$

Evaluating the torque on sides 2 and 4, we find

$$d\mathbf{T}_2 + d\mathbf{T}_4 = dx\,dy\,IB_{0x}\mathbf{a}_y$$

and the total torque is then

$$d\mathbf{T} = I\,dx\,dy(B_{0x}\mathbf{a}_y - B_{0y}\mathbf{a}_x)$$

The quantity within the parentheses may be represented by a cross product,

$$d\mathbf{T} = I\,dx\,dy(\mathbf{a}_z \times \mathbf{B}_0)$$

or

$$\boxed{d\mathbf{T} = I\,d\mathbf{S} \times \mathbf{B}} \tag{15}$$

where $d\mathbf{S}$ is the vector area of the differential current loop and the subscript on $\mathbf{B}_0$ has been dropped.

We now define the product of the loop current and the vector area of the loop as the differential *magnetic dipole moment* $d\mathbf{m}$, with units of A·m^2. Thus

$$\boxed{d\mathbf{m} = I\,d\mathbf{S}} \tag{16}$$

and

$$\boxed{d\mathbf{T} = d\mathbf{m} \times \mathbf{B}} \tag{17}$$

If we extend the results we obtained in Sect. 4.7 for the differential *electric* dipole by determining the torque produced on it by an *electric* field, we see a similar result,

$$d\mathbf{T} = d\mathbf{p} \times \mathbf{E}$$

Equations (15) and (17) are general results which hold for differential loops of any shape, not just rectangular ones. The torque on a circular or triangular loop is also given in terms of the vector surface or the moment by (15) or (17).

Since we selected a differential current loop so that we might assume **B** was constant throughout it, it follows that the torque on a *planar* loop of any size or shape in a *uniform* magnetic field is given by the same expression,

$$\boxed{\mathbf{T} = I\mathbf{S} \times \mathbf{B} = \mathbf{m} \times \mathbf{B}} \tag{18}$$

We should note that the torque on the current loop always tends to turn the loop so as to align the magnetic field produced by the loop with the applied magnetic field that is causing the torque. This is perhaps the easiest way to determine the direction of the torque.

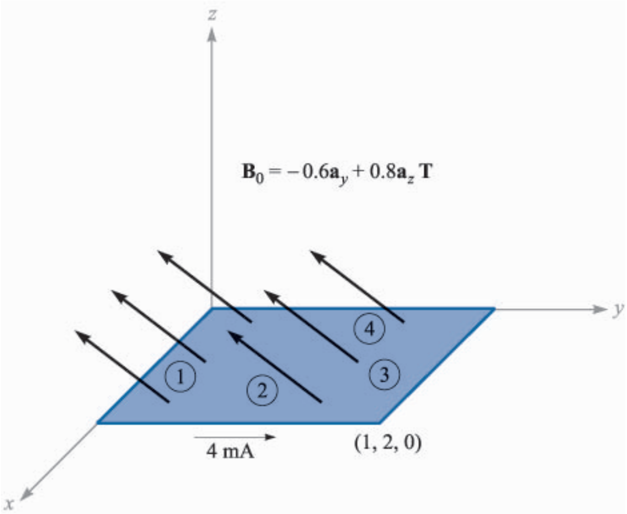


FIGURE 9.7
A rectangular loop is located in a uniform magnetic flux density $\mathbf{B}_0$.

Example 9.3

To illustrate some force and torque calculations, consider the rectangular loop shown in Fig. 9.7. Calculate the torque by using $\mathbf{T} = I\mathbf{S} \times \mathbf{B}$.

Solution. The loop has dimensions of 1 m by 2 m and lies in the uniform field $\mathbf{B}_0 = -0.6\mathbf{a}_y + 0.8\mathbf{a}_z$ T. The loop current is 4 mA, a value that is sufficiently small to avoid causing any magnetic field that might affect $\mathbf{B}_0$.

We have

$$\mathbf{T} = 4 \times 10^{-3}[(1)(2)\mathbf{a}_z] \times (-0.6\mathbf{a}_y + 0.8\mathbf{a}_z) = 4.8\mathbf{a}_x \quad \text{mN} \cdot \text{m}$$

Thus, the loop tends to rotate about an axis parallel to the positive x axis. The small magnetic field produced by the 4-mA loop current tends to line up with $\mathbf{B}_0$.

Example 9.4

Now let us find the torque once more, this time by calculating the total force and torque contribution for each side.

Solution. On side 1 we have

$$\mathbf{F}_1 = I\mathbf{L}_1 \times \mathbf{B}_0 = 4 \times 10^{-3}(1\mathbf{a}_x) \times (-0.6\mathbf{a}_y + 0.8\mathbf{a}_z)$$
$$= -3.2\mathbf{a}_y - 2.4\mathbf{a}_z \quad \text{mN}$$

On side 3 we obtain the negative of this result,

$$\mathbf{F}_3 = 3.2\mathbf{a}_y + 2.4\mathbf{a}_z \quad \text{mN}$$

Next we attack side 2:

$$\mathbf{F}_2 = I\mathbf{L}_2 \times \mathbf{B}_0 = 4 \times 10^{-3}(2\mathbf{a}_y) \times (-0.6\mathbf{a}_y + 0.8\mathbf{a}_z)$$
$$= 6.4\mathbf{a}_x \quad \text{mN}$$

with side 4 again providing the negative of this result,

$$\mathbf{F}_4 = -6.4\mathbf{a}_x \quad \text{mN}$$

Since these forces are distributed uniformly along each of the sides, we treat each force as if it were applied at the center of the side. The origin for the torque may be established anywhere since the sum of the forces is zero, and we choose the center of the loop. Thus,

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \mathbf{T}_4 = \mathbf{R}_1 \times \mathbf{F}_1 + \mathbf{R}_2 \times \mathbf{F}_2 + \mathbf{R}_3 \times \mathbf{F}_3 + \mathbf{R}_4 \times \mathbf{F}_4$$
$$= (-1\mathbf{a}_y) \times (-3.2\mathbf{a}_y - 2.4\mathbf{a}_z) + (0.5\mathbf{a}_x) \times (6.4\mathbf{a}_x)$$
$$+ (1\mathbf{a}_y) \times (3.2\mathbf{a}_y + 2.4\mathbf{a}_z) + (-0.5\mathbf{a}_x) \times (-6.4\mathbf{a}_x)$$
$$= 2.4\mathbf{a}_x + 2.4\mathbf{a}_x = 4.8\mathbf{a}_x \quad \text{mN} \cdot \text{m}$$

Crossing the loop moment with the magnetic flux density is certainly easier.

✔ **D9.5.** A conducting filamentary triangle joins points $A(3, 1, 1)$, $B(5, 4, 2)$, and $C(1, 2, 4)$. The segment AB carries a current of 0.2 A in the $\mathbf{a}_{AB}$ direction. There is present a magnetic field $\mathbf{B} = 0.2\mathbf{a}_x - 0.1\mathbf{a}_y + 0.3\mathbf{a}_z$ T. Find: (*a*) the force on segment BC; (*b*) the force on the triangular loop; (*c*) the torque on the loop about an origin at A; (*d*) the torque on the loop about an origin at C.

Ans. $-0.08\mathbf{a}_x + 0.32\mathbf{a}_y + 016\mathbf{a}_z$ N; 0;$-0.16\mathbf{a}_x - 0.08\mathbf{a}_y + 0.08\mathbf{a}_z$ N·m; $-0.16\mathbf{a}_x - 0.08\mathbf{a}_y + 0.08\mathbf{a}_z$ N·m

9.5 THE NATURE OF MAGNETIC MATERIALS

We are now in a position to combine our knowledge of the action of a magnetic field on a current loop with a simple model of an atom and obtain some appreciation of the difference in behavior of various types of materials in magnetic fields.

Although accurate quantitative results can only be predicted through the use of quantum theory, the simple atomic model which assumes that there is a central positive nucleus surrounded by electrons in various circular orbits yields reasonable quantitative results and provides a satisfactory qualitative theory. An electron in an orbit is analogous to a small current loop (in which the current is directed oppositely to the direction of electron travel) and as such experiences a torque in an external magnetic field, the torque tending to align the magnetic field produced by the orbiting electron with the external magnetic field. If there were no other magnetic moments to consider, we would then conclude that all the orbiting electrons in the material would shift in such a way as to add their magnetic fields to the applied field, and thus that the resultant magnetic field at

any point in the material would be greater than it would be at that point if the material were not present.

A second moment, however, is attributed to *electron spin*. Although it is tempting to model this phenomenon by considering the electron as spinning about its own axis and thus generating a magnetic dipole moment, satisfactory quantitative results are not obtained from such a theory. Instead, it is necessary to digest the mathematics of relativistic quantum theory to show that an electron may have a spin magnetic moment of about $\pm 9 \times 10^{-24}\,\mathrm{A \cdot m^2}$; the plus and minus signs indicate that alignment aiding or opposing an external magnetic field is possible. In an atom with many electrons present, only the spins of those electrons in shells which are not completely filled will contribute to a magnetic moment for the atom.

A third contribution to the moment of an atom is caused by *nuclear spin*. Although this factor provides a negligible effect on the overall magnetic properties of materials, it is the basis of the nuclear magnetic resonance imaging (MRI) procedure now provided by many of the larger hospitals.

Thus each atom contains many different component moments, and their combination determines the magnetic characteristics of the material and provides its general magnetic classification. We shall describe briefly six different types of material: diamagnetic, paramagnetic, ferromagnetic, antiferromagnetic, ferrimagnetic, and superparamagnetic.

Let us first consider those atoms in which the small magnetic fields produced by the motion of the electrons in their orbits and those produced by the electron spin combine to produce a net field of zero. Note that we are considering here the fields produced by the electron motion itself in the absence of any external magnetic field; we might also describe this material as one in which the permanent magnetic moment $\mathbf{m}_0$ of each atom is zero. Such a material is termed *diamagnetic*. It would seem, therefore, that an external magnetic field would produce no torque on the atom, no realignment of the dipole fields, and consequently an internal magnetic field that is the same as the applied field. With an error that only amounts to about one part in a hundred thousand, this is correct.

Let us select an orbiting electron whose moment $\mathbf{m}$ is in the same direction as the applied field $\mathbf{B}_0$ (Fig. 9.8). The magnetic field produces an outward force on the orbiting electron. Since the orbital radius is quantized and cannot change, the inward Coulomb force of attraction is also unchanged. The force unbalance created by the outward magnetic force must therefore be compensated for by a reduced orbital velocity. Hence, the orbital moment decreases, and a smaller internal field results.

If we had selected an atom for which $\mathbf{m}$ and $\mathbf{B}_0$ were opposed, the magnetic force would be inward, the velocity would increase, the orbital moment would increase, and greater cancellation of $\mathbf{B}_0$ would occur. Again a smaller internal field would result.

Metallic bismuth shows a greater diamagnetic effect than most other diamagnetic materials, among which are hydrogen, helium, the other "inert" gases,

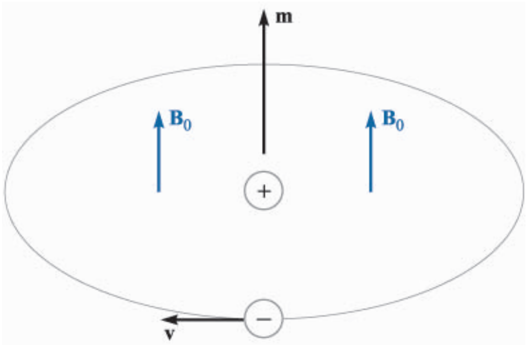


FIGURE 9.8
An orbiting electron is shown having a magnetic moment **m** in the same direction as an applied field $\mathbf{B}_0$.

sodium chloride, copper, gold, silicon, germanium, graphite, and sulfur. We should also realize that the diamagnetic effect is present in all materials, because it arises from an interaction of the external magnetic field with every orbiting electron; however, it is overshadowed by other effects in the materials we shall consider next.

Now let us discuss an atom in which the effects of the electron spin and orbital motion do not quite cancel. The atom as a whole has a small magnetic moment, but the random orientation of the atoms in a larger sample produces an *average* magnetic moment of zero. The material shows no magnetic effects in the absence of an external field. When an external field is applied, however, there is a small torque on each atomic moment, and these moments tend to become aligned with the external field. This alignment acts to increase the value of **B** within the material over the external value. However, the diamagnetic effect is still operating on the orbiting electrons and may counteract the above increase. If the net result is a decrease in **B**, the material is still called diamagnetic. However, if there is an increase in **B**, the material is termed *paramagnetic*. Potassium, oxygen, tungsten, and the rare earth elements and many of their salts, such as erbium chloride, neodymium oxide, and yttrium oxide, one of the materials used in masers, are examples of paramagnetic substances.

The remaining four classes of material, ferromagnetic, antiferromagnetic, ferrimagnetic, and superparamagnetic, all have strong atomic moments. Moreover, the interaction of adjacent atoms causes an alignment of the magnetic moments of the atoms in either an aiding or exactly opposing manner.

In *ferromagnetic* materials each atom has a relatively large dipole moment, caused primarily by uncompensated electron spin moments. Interatomic forces cause these moments to line up in a parallel fashion over regions containing a large number of atoms. These regions are called *domains*, and they may have a variety of shapes and sizes ranging from one micrometer to several centimeters, depending on the size, shape, material, and magnetic history of the sample. Virgin ferromagnetic materials will have domains which each have a strong magnetic moment; the domain moments, however, vary in direction from domain to domain. The overall effect is therefore one of cancellation, and the

material as a whole has no magnetic moment. Upon application of an external magnetic field, however, those domains which have moments in the direction of the applied field increase their size at the expense of their neighbors, and the internal magnetic field increases greatly over that of the external field alone. When the external field is removed, a completely random domain alignment is not usually attained, and a residual, or remnant, dipole field remains in the macroscopic structure. The fact that the magnetic moment of the material is different after the field has been removed, or that the magnetic state of the material is a function of its magnetic history, is called *hysteresis*, a subject which will be discussed again when magnetic circuits are studied a few pages from now.

Ferromagnetic materials are not isotropic in single crystals, and we shall therefore limit our discussion to polycrystalline materials, except for mentioning that one of the characteristics of anisotropic magnetic materials is magnetostriction, or the change in dimensions of the crystal when a magnetic field is impressed on it.

The only elements that are ferromagnetic at room temperature are iron, nickel, and cobalt, and they lose all their ferromagnetic characteristics above a temperature called the Curie temperature, which is 1043 K (770°C) for iron. Some alloys of these metals with each other and with other metals are also ferromagnetic, as for example alnico, an aluminum-nickel-cobalt alloy with a small amount of copper. At lower temperatures some of the rare earth elements, such as gadolinium and dysprosium, are ferromagnetic. It is also interesting that some alloys of nonferromagnetic metals are ferromagnetic, such as bismuth-manganese and copper-manganese-tin.

In *antiferromagnetic* materials, the forces between adjacent atoms cause the atomic moments to line up in an antiparallel fashion. The net magnetic moment is zero, and antiferromagnetic materials are affected only slightly by the presence of an external magnetic field. This effect was first discovered in manganese oxide, but several hundred antiferromagnetic materials have been identified since then. Many oxides, sulfides, and chlorides are included, such as nickel oxide (NiO), ferrous sulfide (FeS), and cobalt chloride ($CoCl_2$). Antiferromagnetism is only present at relatively low temperatures, often well below room temperature. The effect is not of engineering importance at present.

The *ferrimagnetic* substances also show an antiparallel alignment of adjacent atomic moments, but the moments are not equal. A large response to an external magnetic field therefore occurs, although not as large as that in ferromagnetic materials. The most important group of ferrimagnetic materials are the *ferrites*, in which the conductivity is low, several orders of magnitude less than that of semiconductors. The fact that these substances have greater resistance than the ferromagnetic materials results in much smaller induced currents in the material when alternating fields are applied, as for example in transformer cores which operate at the higher frequencies. The reduced currents (eddy currents) lead to lower ohmic losses in the transformer core. The iron oxide magnetite (Fe_3O_4), a nickel-zinc ferrite ($Ni_{1/2}Zn_{1/2}Fe_2O_4$), and a nickel ferrite ($NiFe_2O_4$)

TABLE 9.1
Characteristics of magnetic materials

Classification	Magnetic moments	B values	Comments
Diamagnetic	$\mathbf{m}_{orb} + \mathbf{m}_{spin} = 0$	$B_{int} < B_{appl}$	$B_{int} \doteq B_{appl}$
Paramagnetic	$\mathbf{m}_{orb} + \mathbf{m}_{spin} = \text{small}$	$B_{int} > B_{appl}$	$B_{int} \doteq B_{appl}$
Ferromagnetic	$\lvert\mathbf{m}_{spin}\rvert \gg \lvert\mathbf{m}_{orb}\rvert$	$B_{int} \gg B_{appl}$	Domains
Antiferromagnetic	$\lvert\mathbf{m}_{spin}\rvert \gg \lvert\mathbf{m}_{orb}\rvert$	$B_{int} \doteq B_{appl}$	Adjacent moments oppose
Ferrimagnetic	$\lvert\mathbf{m}_{spin}\rvert > \lvert\mathbf{m}_{orb}\rvert$	$B_{int} > B_{appl}$	Unequal adjacent moments oppose; low σ
Superparamagnetic	$\lvert\mathbf{m}_{spin}\rvert \gg \lvert\mathbf{m}_{orb}\rvert$	$B_{int} > B_{appl}$	Nonmagnetic matrix; recording tapes

are examples of this class of materials. Ferrimagnetism also disappears above the Curie temperature.

Superparamagnetic materials are composed of an assemblage of ferromagnetic particles in a nonferromagnetic matrix. Although domains exist within the individual particles, the domain walls cannot penetrate the intervening matrix material to the adjacent particle. An important example is the magnetic tape used in audiotape or videotape recorders.

Table 9.1 summarizes the characteristics of the six types of magnetic materials discussed above.

9.6 MAGNETIZATION AND PERMEABILITY

To place our description of magnetic materials on a more quantitative basis, we shall now devote a page or so to showing how the magnetic dipoles act as a distributed source for the magnetic field. Our result will be an equation that looks very much like Ampère's circuital law, $\oint \mathbf{H} \cdot d\mathbf{L} = I$. The current, however, will be the movement of *bound* charges (orbital electrons, electron spin, and nuclear spin), and the field, which has the dimensions of $\mathbf{H}$, will be called the magnetization $\mathbf{M}$. The current produced by the bound charges is called a *bound current* or *Amperian current*.

Let us begin by defining the magnetization $\mathbf{M}$ in terms of the magnetic dipole moment $\mathbf{m}$. The bound current I_b circulates about a path enclosing a differential area $d\mathbf{S}$, establishing a dipole moment (A·m^2),

$$\mathbf{m} = I_b d\mathbf{S}$$

If there are n magnetic dipoles per unit volume and we consider a volume Δv, then the total magnetic dipole moment is found by the vector sum

$$\mathbf{m}_{total} = \sum_{i=1}^{n\Delta v} \mathbf{m}_i \tag{19}$$

Each of the $\mathbf{m}_i$ may be different. Next, we define the *magnetization* $\mathbf{M}$ as the *magnetic dipole moment per unit volume*,

$$\mathbf{M} = \lim_{\Delta v \to 0} \frac{1}{\Delta v} \sum_{i=1}^{n\Delta v} \mathbf{m}_i$$

and see that its units must be the same as for **H**, amperes per meter.

Now let us consider the effect of some alignment of the magnetic dipoles as the result of the application of a magnetic field. We shall investigate this alignment along a closed path, a short portion of which is shown in Fig. 9.9. The figure shows several magnetic moments **m** that make an angle θ with the element of path $d\mathbf{L}$; each moment consists of a bound current I_b circulating about an area $d\mathbf{S}$. We are therefore considering a small volume, $dS \cos\theta dL$, or $d\mathbf{S} \cdot d\mathbf{L}$, within which there are $nd\mathbf{S} \cdot d\mathbf{L}$ magnetic dipoles. In changing from a random orientation to this partial alignment, the bound current crossing the surface enclosed by the path (to our left as we travel in the $\mathbf{a}_L$ direction in Fig. 9.9) has increased by I_b for each of the $nd\mathbf{S} \cdot d\mathbf{L}$ dipoles. Thus

$$dI_b = nI_b d\mathbf{S} \cdot d\mathbf{L} = \mathbf{M} \cdot d\mathbf{L} \tag{20}$$

and within an entire closed contour,

$$I_b = \oint \mathbf{M} \cdot d\mathbf{L} \tag{21}$$

Equation (21) merely says that if we go around a closed path and find dipole moments going our way more often than not, there will be a corresponding current composed of, for example, orbiting electrons crossing the interior surface.

This last expression has some resemblance to Ampère's circuital law, and we may now generalize the relationship between **B** and **H** so that it applies to media other than free space. Our present discussion is based on the forces and torques on differential current loops in a **B** field, and we therefore take **B** as our fundamental quantity and seek an improved definition of **H**. We thus write Ampère's circuital law in terms of the *total* current, bound plus free,

$$\oint \frac{\mathbf{B}}{\mu_0} \cdot d\mathbf{L} = I_T \tag{22}$$

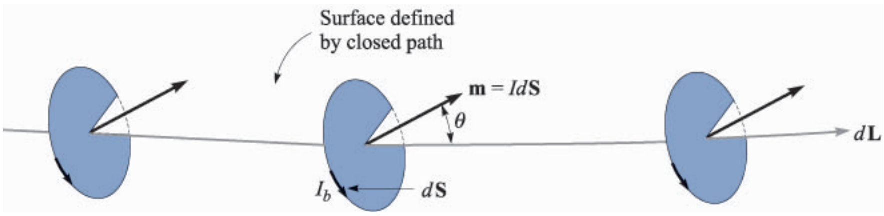


FIGURE 9.9
A section $d\mathbf{L}$ of a closed path along which magnetic dipoles have been partially aligned by some external magnetic field. The alignment has caused the bound current crossing the surface defined by the closed path to increase by $nI_b d\mathbf{S} \cdot d\mathbf{L}$ amperes.

where

$$I_T = I_b + I$$

and I is the total *free* current enclosed by the closed path. Note that the free current appears without subscript since it is the most important type of current and will be the only current appearing in Maxwell's equations.

Combining these last three equations, we obtain an expression for the free current enclosed,

$$I = I_T - I_b = \oint \left(\frac{\mathbf{B}}{\mu_0} - \mathbf{M} \right) \cdot d\mathbf{L} \tag{23}$$

We may now define $\mathbf{H}$ in terms of $\mathbf{B}$ and $\mathbf{M}$,

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M} \tag{24}$$

and we see that $\mathbf{B} = \mu_0 \mathbf{H}$ in free space where the magnetization is zero. This relationship is usually written in a form that avoids fractions and minus signs:

$$\boxed{\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})} \tag{25}$$

We may now use our newly defined $\mathbf{H}$ field in (23),

$$I = \oint \mathbf{H} \cdot d\mathbf{L} \tag{26}$$

obtaining Ampère's circuital law in terms of the free currents.

Utilizing the several current densities, we have

$$I_b = \oint_S \mathbf{J}_b \cdot d\mathbf{S}$$

$$I_T = \oint_S \mathbf{J}_T \cdot d\mathbf{S}$$

$$I = \oint_S \mathbf{J} \cdot d\mathbf{S}$$

With the help of Stokes' theorem, we may therefore transform (21), (26), and (22) into the equivalent curl relationships:

$$\nabla \times \mathbf{M} = \mathbf{J}_b$$

$$\nabla \times \frac{\mathbf{B}}{\mu_0} = \mathbf{J}_T$$

$$\boxed{\nabla \times \mathbf{H} = \mathbf{J}} \tag{27}$$

We shall emphasize only (26) and (27), the two expressions involving the free charge, in the work that follows.

The relationship between **B**, **H**, and **M** expressed by (25) may be simplified for linear isotropic media where a magnetic susceptibility χ_m can be defined:

$$\boxed{\mathbf{M} = \chi_m \mathbf{H}} \tag{28}$$

Thus we have

$$\mathbf{B} = \mu_0(\mathbf{H} + \chi_m \mathbf{H})$$
$$= \mu_0 \mu_R \mathbf{H}$$

where

$$\mu_R = 1 + \chi_m \tag{29}$$

is defined as the *relative permeability* μ_R. We next define the *permeability* μ:

$$\mu = \mu_0 \mu_R \tag{30}$$

and this enables us to write the simple relationship between **B** and **H**,

$$\mathbf{B} = \mu \mathbf{H} \tag{31}$$

Example 9.5

Given a ferrite material which we shall specify to be operating in a linear mode with $B = 0.05\,\text{T}$, let us assume $\mu_R = 50$, and calculate values for χ_m, M, and H.

Solution. Since $\mu_R = 1 + \chi_m$, we have

$$\chi_m = \mu_R - 1 = 49$$

Also,

$$B = \mu_R \mu_0 H$$

and

$$H = \frac{0.05}{50 \times 4\pi \times 10^{-7}} = 796 \quad \text{A/m}$$

The magnetization is $\chi_m H$, or 39 000 A/m. The alternate ways of relating B and H are, first,

$$B = \mu_0(H + M)$$

or

$$0.05 = 4\pi \times 10^{-7}(796 + 39\,000)$$

showing that Amperian currents produce 49 times the magnetic field intensity that the free charges do; and second,

$$B = \mu_R \mu_0 H$$

or

$$0.05 = 50 \times 4\pi \times 10^{-7} \times 796$$

where we utilize a relative permeability of 50 and let this quantity account completely for the notion of the bound charges. We shall emphasize the latter interpretation in the chapters that follow.

The first two laws that we investigated for magnetic fields were the Biot-Savart law and Ampère's circuital law. Both were restricted to free space in their application. We may now extend their use to any homogeneous, linear, isotropic magnetic material that may be described in terms of a relative permeability μ_R.

Just as we found for anisotropic dielectric materials, the permeability of an anisotropic magnetic material must be given as a 3×3 matrix, while **B** and **H** are both 3×1 matrices. We have

$$B_x = \mu_{xx} H_x + \mu_{xy} H_y + \mu_{xz} H_z$$
$$B_y = \mu_{yx} H_x + \mu_{yy} H_y + \mu_{yz} H_z$$
$$B_z = \mu_{zx} H_x + \mu_{zy} H_y + \mu_{zz} H_z$$

For anisotropic materials, then, $\mathbf{B} = \mu \mathbf{H}$ is a matrix equation; however $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ remains valid, although **B**, **H**, and **M** are no longer parallel in general. The most common anisotropic magnetic material is a single ferromagnetic crystal, although thin magnetic films also exhibit anisotropy. Most applications of ferromagnetic materials, however, involve polycrystalline arrays that are much easier to make.

Our definitions of susceptibility and permeability also depend on the assumption of linearity. Unfortunately, this is true only in the less interesting paramagnetic and diamagnetic materials for which the relative permeability rarely differs from unity by more than one part in a thousand. Some typical values of the susceptibility for diamagnetic materials are hydrogen, -2×10^{-5}; copper, -0.9×10^{-5}; germanium, -0.8×10^{-5}; silicon, -0.3×10^{-5}; and graphite, -12×10^{-5}. Several representative paramagnetic susceptibilities are oxygen, 2×10^{-6}; tungsten, 6.8×10^{-5}; ferric oxide (Fe_2O_3), 1.4×10^{-3}; and yttrium oxide (Y_2O_3), 0.53×10^{-6}. If we simply take the ratio of B to $\mu_0 H$ as the relative permeability of a ferromagnetic material, typical values of μ_R would range from 10 to 100 000. Diamagnetic, paramagnetic, and antiferromagnetic materials are commonly said to be nonmagnetic.

✔ **D9.6.** Find the magnetization in a magnetic material where: (*a*) $\mu = 1.8 \times 10^{-5}$ H/m and $H = 120$ A/m; (*b*) $\mu_R = 22$, there are 8.3×10^{28} atoms/m^3, and each atom has a dipole moment of 4.5×10^{-27} A·m^2; (*c*) $B = 300\,\mu$T and $\chi_m = 15$.

Ans. 1599 A/m; 374 A/m; 224 A/m

✔ **D9.7.** The magnetization in a magnetic material for which $\chi_m = 8$ is given in a certain region as $150z^2\mathbf{a}_x$ A/m. At $z = 4$ cm, find the magnitude of: (*a*) $\mathbf{J}_T$; (*b*) $\mathbf{J}$; (*c*) $\mathbf{J}_b$.

Ans. 108 A/m^2; 12 A/m^2; 96 A/m^2

9.7 MAGNETIC BOUNDARY CONDITIONS

We should have no difficulty in arriving at the proper boundary conditions to apply to **B**, **H**, and **M** at the interface between two different magnetic materials, for we have solved similar problems for both conducting materials and dielectrics. We need no new techniques.

Fig. 9.10 shows a boundary between two isotropic homogeneous linear materials with permeabilities μ_1 and μ_2. The boundary condition on the normal components is determined by allowing the surface to cut a small cylindrical gaussian surface. Applying Gauss's law for the magnetic field from Sec. 8.5,

$$\oint_S \mathbf{B} \cdot d\mathbf{S} = 0$$

we find that

$$B_{N1}\Delta S - B_{N2}\Delta S = 0$$

or

$$\boxed{B_{N2} = B_{N1}} \tag{32}$$

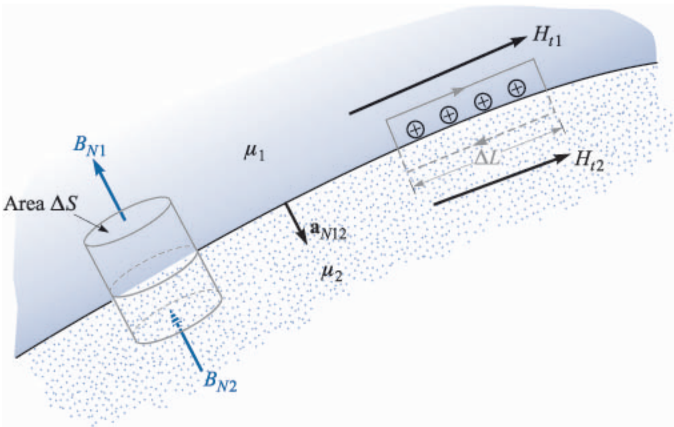


FIGURE 9.10
A gaussian surface and a closed path are constructed at the boundary between media 1 and 2, having permeabilities of μ_1 and μ_2, respectively. From this we determine the boundary conditions $B_{N1} = B_{N2}$ and $H_{t1} - H_{t2} = K$, the component of the surface current density directed into the page.

Thus

$$H_{N2} = \frac{\mu_1}{\mu_2} H_{N1} \tag{33}$$

The normal component of **B** is continuous, but the normal component of **H** is discontinuous by the ratio μ_1/μ_2.

The relationship between the normal components of **M**, of course, is fixed once the relationship between the normal components of **H** is known. For linear magnetic materials, the result is written simply as

$$M_{N2} = \chi_{m2} \frac{\mu_1}{\mu_2} H_{N1} = \frac{\chi_{m2}\mu_1}{\chi_{m1}\mu_2} M_{N1} \tag{34}$$

Next, Ampère's circuital law

$$\oint \mathbf{H} \cdot d\mathbf{L} = I$$

is applied about a small closed path in a plane normal to the boundary surface, as shown to the right in Fig. 9.10. Taking a clockwise trip around the path, we find that

$$H_{t1}\Delta L - H_{t2}\Delta L = K\Delta L$$

where we assume that the boundary may carry a surface current **K** whose component normal to the plane of the closed path is K. Thus

$$\boxed{H_{t1} - H_{t2} = K} \tag{35}$$

The directions are specified more exactly by using the cross product to identify the tangential components,

$$(\mathbf{H}_1 - \mathbf{H}_2) \times \mathbf{a}_{N12} = \mathbf{K}$$

where $\mathbf{a}_{N12}$ is the unit normal at the boundary directed from region 1 to region 2. An equivalent formulation in terms of the vector tangential components may be more convenient for **H**:

$$\mathbf{H}_{t1} - \mathbf{H}_{t2} = \mathbf{a}_{N12} \times \mathbf{K}$$

For tangential **B**, we have

$$\frac{B_{t1}}{\mu_1} - \frac{B_{t2}}{\mu_2} = K \tag{36}$$

The boundary condition on the tangential component of the magnetization for linear materials is therefore

$$M_{t2} = \frac{\chi_{m2}}{\chi_{m1}} M_{t1} - \chi_{m2} K \tag{37}$$

The last three boundary conditions on the tangential components are much simpler, of course, if the surface current density is zero. This is a free current density, and it must be zero if neither material is a conductor.

Example 9.6

To illustrate these relationships with an example, let us assume that $\mu = \mu_1 = 4\,\mu\text{H/m}$ in region 1 where $z > 0$, while $\mu_2 = 7\,\mu\text{H/m}$ wherever $z < 0$. Moreover, let $\mathbf{K} = 80\mathbf{a}_x$ A/m on the surface $z = 0$. We establish a field, $\mathbf{B}_1 = 2\mathbf{a}_x - 3\mathbf{a}_y + \mathbf{a}_z$ mT, in region 1 and seek the value of $\mathbf{B}_2$.

Solution. The normal component of $\mathbf{B}_1$ is

$$\mathbf{B}_{N1} = (\mathbf{B}_1 \cdot \mathbf{a}_{N12})\mathbf{a}_{N12} = [(2\mathbf{a}_x - 3\mathbf{a}_y + \mathbf{a}_z)\cdot(-\mathbf{a}_z)](-\mathbf{a}_z) = \mathbf{a}_z \quad \text{mT}$$

Thus,

$$\mathbf{B}_{N2} = \mathbf{B}_{N1} = \mathbf{a}_z \quad \text{mT}$$

We next determine the tangential components:

$$\mathbf{B}_{t1} = \mathbf{B}_1 - \mathbf{B}_{N1} = 2\mathbf{a}_x - 3\mathbf{a}_y \quad \text{mT}$$

and

$$\mathbf{H}_{t1} = \frac{\mathbf{B}_{t1}}{\mu_1} = \frac{(2\mathbf{a}_x - 3\mathbf{a}_y)10^{-3}}{4\times 10^{-6}} = 500\mathbf{a}_x - 750\mathbf{a}_y \quad \text{A/m}$$

Thus,

$$\mathbf{H}_{t2} = \mathbf{H}_{t1} - \mathbf{a}_{N12}\times\mathbf{K} = 500\mathbf{a}_x - 750\mathbf{a}_y - (-\mathbf{a}_z)\times 80\mathbf{a}_x$$
$$= 500\mathbf{a}_x - 750\mathbf{a}_y + 80\mathbf{a}_y = 500\mathbf{a}_x - 670\mathbf{a}_y \quad \text{A/m}$$

and

$$\mathbf{B}_{t2} = \mu_2\mathbf{H}_{t2} = 7\times 10^{-6}(500\mathbf{a}_x - 670\mathbf{a}_y) = 3.5\mathbf{a}_x - 4.69\mathbf{a}_y \quad \text{mT}$$

Therefore,

$$\mathbf{B}_2 = \mathbf{B}_{N2} + \mathbf{B}_{t2} = 3.5\mathbf{a}_x - 4.69\mathbf{a}_y + \mathbf{a}_z \quad \text{mT}$$

✔ **D9.8.** Let the permittivity be $5\,\mu\text{H/m}$ in region A where $x < 0$, and $20\,\mu\text{H/m}$ in region B where $x > 0$. If there is a surface current density $\mathbf{K} = 150\mathbf{a}_y - 200\mathbf{a}_z$ A/m at $x = 0$, and if $H_A = 300\mathbf{a}_x - 400\mathbf{a}_y + 500\mathbf{a}_z$ A/m, find: (*a*) $|\mathbf{H}_{tA}|$; (*b*) $|\mathbf{H}_{NA}|$; (*c*) $|\mathbf{H}_{tB}|$; (*d*) $|\mathbf{H}_{NB}|$.

Ans. 640 A/m; 300 A/m; 695 A/m; 75 A/m

9.8 THE MAGNETIC CIRCUIT

In this section we shall digress briefly to discuss the fundamental techniques involved in solving a class of magnetic problems known as magnetic circuits. As we shall see shortly, the name arises from the great similarity to the dc-resistive-circuit analysis with which it is assumed we are all facile. The only

important difference lies in the nonlinear nature of the ferromagnetic portions of the magnetic circuit; the methods which must be adopted are similar to those required in nonlinear electric circuits which contain diodes, thermistors, incandescent filaments, and other nonlinear elements.

As a convenient starting point, let us identify those field equations upon which resistive circuit analysis is based. At the same time we shall point out or derive the analogous equations for the magnetic circuit. We begin with the electrostatic potential and its relationship to electric field intensity,

$$\mathbf{E} = -\nabla V \tag{38a}$$

The scalar magnetic potential has already been defined, and its analogous relation to the magnetic field intensity is

$$\boxed{\mathbf{H} = -\nabla V_m} \tag{38b}$$

In dealing with magnetic circuits, it is convenient to call V_m the *magnetomotive force*, or mmf, and we shall acknowledge the analogy to the electromotive force, or emf, by doing so. The units of the mmf are, of course, amperes, but it is customary to recognize that coils with many turns are often employed by using the term "ampere-turns." Remember that no current may flow in any region in which V_m is defined.

The electric potential difference between points A and B may be written as

$$V_{AB} = \int_A^B \mathbf{E} \cdot d\mathbf{L} \tag{39a}$$

and the corresponding relationship between the mmf and the magnetic field intensity,

$$\boxed{V_{mAB} = \int_A^B \mathbf{H} \cdot d\mathbf{L}} \tag{39b}$$

was developed in Chap. 8, where we learned that the path selected must not cross the chosen barrier surface.

Ohm's law for the electric circuit has the point form

$$\mathbf{J} = \sigma \mathbf{E} \tag{40a}$$

and we see that the magnetic flux density will be the analog of the current density,

$$\boxed{\mathbf{B} = \mu \mathbf{H}} \tag{40b}$$

To find the total current, we must integrate:

$$I = \int_S \mathbf{J} \cdot d\mathbf{S} \tag{41a}$$

A corresponding operation is necessary to determine the total magnetic flux flowing through the cross section of a magnetic circuit:

$$\Phi = \int_S \mathbf{B} \cdot d\mathbf{S} \tag{41b}$$

We then defined resistance as the ratio of potential difference and current, or

$$V = IR \tag{42a}$$

and we shall now define *reluctance* as the ratio of the magnetomotive force to the total flux; thus

$$V_m = \Phi\Re \tag{42b}$$

where reluctance is measured in ampere-turns per weber (A·t/Wb). In resistors which are made of a linear isotropic homogeneous material of conductivity σ and have a uniform cross section of area S and length d, the total resistance is

$$R = \frac{d}{\sigma S} \tag{43a}$$

If we are fortunate enough to have such a linear isotropic homogeneous magnetic material of length d and uniform cross section S, then the total reluctance is

$$\Re = \frac{d}{\mu S} \tag{43b}$$

The only such material to which we shall commonly apply this relationship is air.

Finally, let us consider the analog of the source voltage in an electric circuit. We know that the closed line integral of $\mathbf{E}$ is zero,

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0$$

In other words, Kirchhoff's voltage law states that the rise in potential through the source is exactly equal to the fall in potential through the load. The expression for magnetic phenomena takes on a slightly different form,

$$\oint \mathbf{H} \cdot d\mathbf{L} = I_{total}$$

for the closed line integral is not zero. Since the total current linked by the path is usually obtained by allowing a current I to flow through an N-turn coil, we may express this result as

$$\oint \mathbf{H} \cdot d\mathbf{L} = NI \tag{44}$$

In an electric circuit the voltage source is a part of the closed path; in the magnetic circuit the current-carrying coil will surround or link the magnetic circuit. In tracing a magnetic circuit, we shall not be able to identify a pair of terminals at which the magnetomotive force is applied. The analogy is closer here to a pair of coupled circuits in which induced voltages exist (and in which we shall see in Chap. 10 that the closed line integral of $\mathbf{E}$ is also not zero).

Let us try out some of these ideas on a simple magnetic circuit. In order to avoid the complications of ferromagnetic materials at this time, we shall assume that we have an air-core toroid with 500 turns, a cross-sectional area of 6 cm^2, a mean radius of 15 cm, and a coil current of 4 A. As we already know, the magnetic field is confined to the interior of the toroid, and if we consider the closed path of our magnetic circuit along the mean radius, we link 2000 A·t,

$$V_{m,\,\text{source}} = 2000 \quad \text{A}\cdot\text{t}$$

Although the field in the toroid is not quite uniform, we may assume that it is for all practical purposes and calculate the total reluctance of the circuit as

$$\mathfrak{R} = \frac{d}{\mu S} = \frac{2\pi(0.15)}{4\pi 10^{-7} \times 6 \times 10^{-4}} = 1.25 \times 10^9 \quad \text{A}\cdot\text{t/Wb}$$

Thus

$$\Phi = \frac{V_{m,S}}{\mathfrak{R}} = \frac{2000}{1.25 \times 10^9} = 1.6 \times 10^{-6} \quad \text{Wb}$$

This value of the total flux is in error by less than $\frac{1}{4}$ percent, in comparison with the value obtained when the exact distribution of flux over the cross section is used.

Hence

$$B = \frac{\Phi}{S} = \frac{1.6 \times 10^{-6}}{6 \times 10^{-4}} = 2.67 \times 10^{-3} \quad \text{T}$$

and finally,

$$H = \frac{B}{\mu} = \frac{2.67 \times 10^{-3}}{4\pi 10^{-7}} = 2120 \quad \text{A}\cdot\text{t/m}$$

As a check, we may apply Ampère's circuital law directly in this symmetrical problem,

$$H_\phi 2\pi r = NI$$

and obtain

$$H_\phi = \frac{NI}{2\pi r} = \frac{500 \times 4}{6.28 \times 0.15} = 2120 \quad \text{A/m}$$

at the mean radius.

Our magnetic circuit in this example does not give us any opportunity to find the mmf across different elements in the circuit, for there is only one type of material. The analogous electric circuit is, of course, a single source and a single resistor. We could make it look just as long as the above analysis, however, if we found the current density, the electric field intensity, the total current, the resistance, and the source voltage.

More interesting and more practical problems arise when ferromagnetic materials are present in the circuit. Let us begin by considering the relationship between B and H in such a material. We may assume that we are establishing a curve of B versus H for a sample of ferromagnetic material which is completely demagnetized; both B and H are zero. As we begin to apply an mmf, the flux density also rises, but not linearly, as the experimental data of Fig. 9.11 show near the origin. After H reaches a value of about 100 A·t/m, the flux density rises more slowly and begins to saturate when H is several hundred A·t/m. Having reached partial saturation, let us now turn to Fig. 9.12, where we may continue our experiment at point x by reducing H. As we do so, the effects of hysteresis begin to show, and we do not retrace our original curve. Even after H is zero, $B = B_r$, the remnant flux density. As H is reversed, then brought back to zero, and the complete cycle traced several times, the hysteresis loop of Fig. 9.12 is obtained. The mmf required to reduce the flux density to zero is identified as H_c,

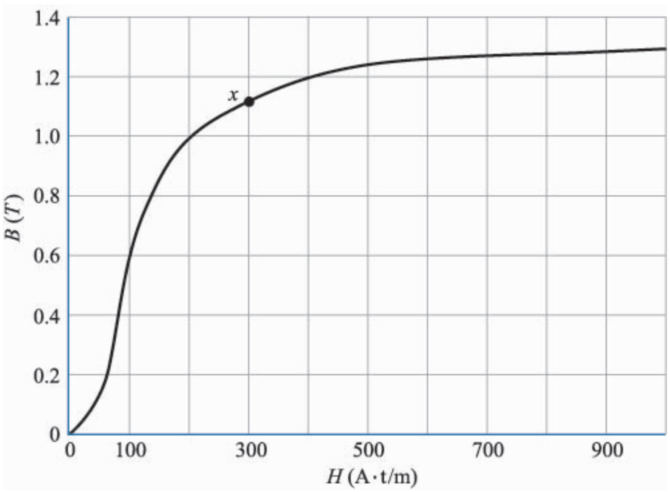


FIGURE 9.11
Magnetization curve of a sample of silicon sheet steel.

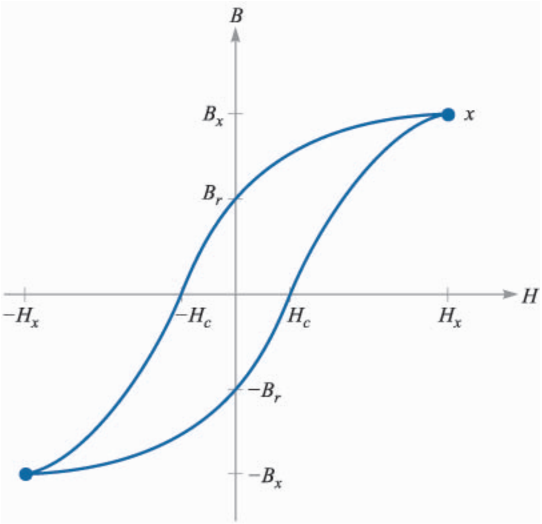


FIGURE 9.12
A hysteresis loop for silicon steel. The coercive force H_c and remnant flux density B_r are indicated.

the coercive "force." For smaller maximum values of H smaller hysteresis loops are obtained and the locus of the tips is about the same as the virgin magnetization curve of Figure 9.11.

Example 9.7

Let us make use of the magnetization curve for silicon steel to solve a magnetic circuit problem that is slightly different from our previous example. We shall use a steel core in the toroid, except for an air gap of 2 mm. Magnetic circuits with air gaps occur because gaps are deliberately introduced in some devices, such as inductors, which must carry large direct currents, because they are unavoidable in other devices such as rotating machines, or because of unavoidable problems in assembly. There are still 500 turns about the toroid, and we ask what current is required to establish a flux density of 1 T everywhere in the core.

Solution. This magnetic circuit is analogous to an electric circuit containing a voltage source and two resistors, one of which is nonlinear. Since we are given the "current," it is easy to find the "voltage" across each series element, and hence the total "emf." In the air gap,

$$\Re_{\text{air}} = \frac{d_{\text{air}}}{\mu S} = \frac{2 \times 10^{-3}}{4\pi 10^{-7} \times 6 \times 10^{-4}} = 2.65 \times 10^6 \quad \text{A}\cdot\text{t/Wb}$$

Knowing the total flux,

$$\Phi = BS = 1(6 \times 10^{-4}) = 6 \times 10^{-4} \quad \text{Wb}$$

which is the same in both steel and air, we may find the mmf required for the gap,

$$V_{m,\text{air}} = (6 \times 10^{-4})(2.65 \times 10^6) = 1590 \quad \text{A}\cdot\text{t}$$

Referring to Fig. 9.11, a magnetic field strength of 200 A·t/m is required to produce a flux density of 1 T in the steel. Thus

$$H_{\text{steel}} = 200 \quad \text{A}\cdot\text{t}$$
$$V_{m,\text{steel}} = H_{\text{steel}} d_{\text{steel}} = 200 \times 0.30\pi$$
$$= 188\ \text{A}\cdot\text{t}$$

The total mmf is therefore 1778 A·t, and a coil current of 3.56 A is required.

We should realize that we have made several approximations in obtaining this answer. We have already mentioned the lack of a completely uniform cross section, or cylindrical symmetry; the path of every flux line is not of the same length. The choice of a "mean" path length can help compensate for this error in problems in which it may be more important than it is in our example. Fringing flux in the air gap is another source of error, and formulas are available by which we may calculate an effective length and cross-sectional area for the gap which will yield more accurate results. There is also a leakage flux between the turns of wire, and in devices containing coils concentrated on one section of the core, a few flux lines bridge the interior of the toroid. Fringing and leakage are problems which seldom arise in the electric circuit because the ratio of the conductivities of air and the conductive or resistive materials used is so high. In contrast, the magnetization curve for silicon steel shows that the ratio of H to B in the steel is about 200 up to the "knee" of the magnetization curve; this compares with a ratio in air of about 800 000. Thus, although flux prefers steel to air by the commanding ratio of 4000 to 1, this is not very close to the ratio of conductivities of, say, 10^{15} for a good conductor and a fair insulator.

Example 9.8

As a last example, let us consider the reverse problem. Given a coil current of 4 A in the previous magnetic circuit, what will the flux density be?

Solution. First let us try to linearize the magnetization curve by a straight line from the origin to $B = 1$, $H = 200$. We then have $B = H/200$ in steel and $B = \mu_0 H$ in air. The two reluctances are found to be 0.314×10^6 for the steel path and 2.65×10^6 for the air gap, or 2.96×10^6 A·t/Wb total. Since V_m is 2000 A·t, the flux is 6.76×10^{-4} Wb, and $B = 1.13$ T. A more accurate solution may be obtained by assuming several values of B and calculating the necessary mmf. Plotting the results enables us to determine the true value of B by interpolation. With this method we obtain $B = 1.10$ T. The good accuracy of the linear model results from the fact that the reluctance of the air gap in a magnetic circuit is often much greater than the reluctance of the ferromagnetic portion of the circuit. A relatively poor approximation for the iron or steel can thus be tolerated.

✔ **D9.9.** Given the magnetic circuit of Fig. 9.13, assume $B = 0.6$ T at the midpoint of the left leg and find: (*a*) $V_{m,\text{air}}$; (*b*) $V_{m,\text{steel}}$; (*c*) the current required in a 1300-turn coil linking the left leg.

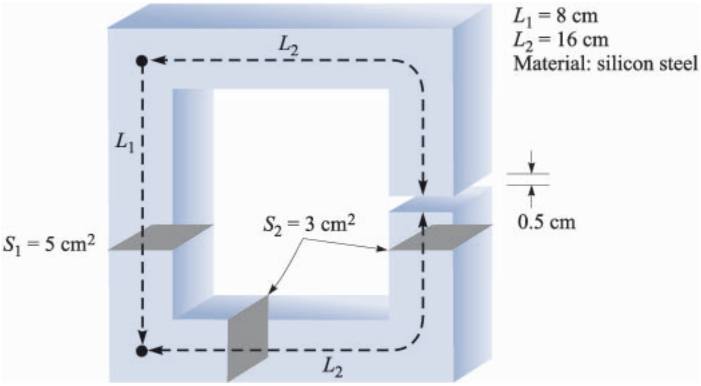


FIGURE 9.13
See Prob. D9.9.

Ans. 3980 A·t; 72 A·t; 3.12 A

✔ **D9.10.** The magnetization curve for material X under normal operating conditions may be approximated by the expression $B = (H/160)(0.25 + e^{-H/320})$, where H is in A/m and B is in T. If a magnetic circuit contains a 12-cm length of material X, as well as a 0.25-mm air gap, assume a uniform cross section of $2.5\,\text{cm}^2$ and find the total mmf required to produce a flux of: (*a*) $10\,\mu$Wb; (*b*) $100\,\mu$Wb.

Ans. 13.14 A·t; 86.7 A·t

9.9 POTENTIAL ENERGY AND FORCES ON MAGNETIC MATERIALS

In the electrostatic field we first introduced the point charge and the experimental law of force between point charges. After defining electric field intensity, electric flux density, and electric potential, we were able to find an expression for the energy in an electrostatic field by establishing the work necessary to bring the prerequisite point charges from infinity to their final resting places. The general expression for energy is

$$W_E = \frac{1}{2}\int_{\text{vol}} \mathbf{D}\cdot\mathbf{E}\,dv \tag{45}$$

where a linear relationship between **D** and **E** is assumed.

This is not as easily done for the steady magnetic field. It would seem that we might assume two simple sources, perhaps two current sheets, find the force on one due to the other, move the sheet a differential distance against this force, and equate the necessary work to the change in energy. If we did, we would be wrong, because Faraday's law (coming up in the next chapter) shows that there will be a voltage induced in the moving current sheet against which the current

must be maintained. Whatever source is supplying the current sheet turns out to receive half the energy we are putting into the circuit by moving it.

In other words, energy density in the magnetic field may be determined more easily after time-varying fields are discussed. We shall develop the appropriate expression in discussing Poynting's theorem in Chap. 11.

An alternate approach would be possible at this time, however, for we might define a magnetostatic field based on assumed magnetic poles (or "magnetic charges"). Using the scalar magnetic potential, we could then develop an energy expression by methods similar to those used in obtaining the electrostatic energy relationship. These new magnetostatic quantities we would have to introduce would be too great a price to pay for one simple result, and we shall therefore merely present the result at this time and show that the same expression arises in the Poynting theorem later. The total energy stored in a steady magnetic field in which **B** is linearly related to **H** is

$$W_H = \frac{1}{2}\int_{\text{vol}} \mathbf{B}\cdot\mathbf{H}\,dv \tag{46}$$

Letting $\mathbf{B} = \mu\mathbf{H}$, we have the equivalent formulations

$$W_H = \frac{1}{2}\int_{\text{vol}} \mu H^2 dv \tag{47}$$

or

$$W_H = \frac{1}{2}\int_{\text{vol}} \frac{B^2}{\mu} dv \tag{48}$$

It is again convenient to think of this energy as being distributed throughout the volume with an energy density of $\frac{1}{2}\mathbf{B}\cdot\mathbf{H}\,\text{J/m}^3$, although we have no mathematical justification for such a statement.

In spite of the fact that these results are valid only for linear media, we may use them to calculate the forces on nonlinear magnetic materials if we focus our attention on the linear media (usually air) which may surround them. For example, suppose that we have a long solenoid with a silicon-steel core. A coil containing n turns/m with a current I surrounds it. The magnetic field intensity in the core is therefore nI A·t/m, and the magnetic flux density can be obtained from the magnetization curve for silicon steel. Let us call this value B_{st}. Suppose that the core is composed of two semi-infinite cylinders[2] which are just touching. We now apply a mechanical force to separate these two sections of the core while keeping the flux density constant. We apply a force F over a distance dL, thus doing work $F\,dL$. Faraday's law does not apply here, for the fields in the core have not changed, and we can therefore use the principle of virtual work to

[2] A semi-infinite cylinder is a cylinder of infinite length having one end located in finite space.

determine that the work we have done in moving one core appears as stored energy in the air gap we have created. By (48) above, this increase is

$$dW_H = F\,dL = \frac{1}{2}\frac{B_{st}^2}{\mu_0}S\,dL$$

where S is the core cross-sectional area. Thus

$$F = \frac{B_{st}^2 S}{2\mu_0}$$

If, for example, the magnetic field intensity is sufficient to produce saturation in the steel, approximately 1.4 T, the force is

$$F = 7.80 \times 10^5 S \quad \text{N}$$

or about 113 lb_f/in^2.

✔ **D9.11.** (*a*) What force is being exerted on the pole faces of the circuit described in Prob. D9.9 and Figure 9.13? (*b*) Is the force trying to open or close the air gap?

Ans. 1194 N; as Wilhelm Eduard Weber would put it, "schliessen"

9.10 INDUCTANCE AND MUTUAL INDUCTANCE

Inductance is the last of the three familiar parameters from circuit theory which we are defining in more general terms. Resistance was defined in Chap. 5 as the ratio of the potential difference between two equipotential surfaces of a conducting material to the total current crossing either equipotential surface. The resistance is a function of conductor geometry and conductivity only. Capacitance was defined in the same chapter as the ratio of the total charge on either of two equipotential conducting surfaces to the potential difference between the surfaces. Capacitance is a function only of the geometry of the two conducting surfaces and the permittivity of the dielectric medium between or surrounding them. The interpretation of resistance and capacitance as circuit elements will be inspected more closely in Sec. 13.1.

As a prelude to defining inductance, we first need to introduce the concept of flux linkage. Let us consider a toroid of N turns in which a current I produces a total flux Φ. We shall assume first that this flux links or encircles each of the N turns, and we also see that each of the N turns links the total flux Φ. The *flux linkage* $N\Phi$ is defined as the product of the number of turns N and the flux Φ linking each of them.[3] For a coil having a single turn, the flux linkage is equal to the total flux.

[3] The symbol λ is commonly used for flux linkages. We shall only occasionally make use of this concept, however, and we will continue to write it as $N\Phi$.

We now define *inductance* (or self-inductance) as the ratio of the total flux linkages to the current which they link,

$$L = \frac{N\Phi}{I} \tag{49}$$

The current I flowing in the N-turn coil produces the total flux Φ and $N\Phi$ flux linkages, where we assume for the moment that the flux Φ links each turn. This definition is applicable only to magnetic media which are linear, so that the flux is proportional to the current. If ferromagnetic materials are present, there is no single definition of inductance which is useful in all cases, and we shall restrict our attention to linear materials.

The unit of inductance is the henry (H), equivalent to one weber-turn per ampere.

Let us apply (49) in a straightforward way to calculate the inductance per meter length of a coaxial cable of inner radius a and outer radius b. We may take the expression for total flux developed as Eq. (42) in Chap. 8,

$$\Phi = \frac{\mu_0 I d}{2\pi} \ln \frac{b}{a}$$

and obtain the inductance rapidly for a length d,

$$L = \frac{\mu_0 d}{2\pi} \ln \frac{b}{a} \quad \text{H}$$

or, on a per-meter basis,

$$L = \frac{\mu_0}{2\pi} \ln \frac{b}{a} \quad \text{H/m} \tag{50}$$

In this case, $N = 1$ turn, and all the flux links all the current.

In the problem of a toroidal coil of N turns and a current I, as shown in Fig. 8.12b, we have

$$B_\phi = \frac{\mu_0 N I}{2\pi\rho}$$

If the dimensions of the cross section are small compared with the mean radius of the toroid ρ_0 then the total flux is

$$\Phi = \frac{\mu_0 N I S}{2\pi\rho_0}$$

where S is the cross-sectional area. Multiplying the total flux by N, we have the flux linkages, and dividing by I, we have the inductance

$$L = \frac{\mu_0 N^2 S}{2\pi\rho_0} \tag{51}$$

Once again we have assumed that all the flux links all the turns, and this is a good assumption for a toroidal coil of many turns packed closely together. Suppose, however, that our toroid has an appreciable spacing between turns, a short part of which might look like Fig. 9.14. The flux linkages are no longer the product of the flux at the mean radius times the total number of turns. In order to obtain the total flux linkages we must look at the coil on a turn-by-turn basis.

$$(N\Phi)_{\text{total}} = \Phi_1 + \Phi_2 + \cdots + \Phi_i + \cdots + \Phi_N$$
$$= \sum_{i=1}^{N} \Phi_i$$

where Φ_i is the flux linking the ith turn. Rather than doing this, we usually rely on experience and empirical quantities called winding factors and pitch factors to adjust the basic formula to apply to the real physical world.

An equivalent definition for inductance may be made using an energy point of view,

$$\boxed{L = \frac{2W_H}{I^2}} \tag{52}$$

where I is the total current flowing in the closed path and W_H is the energy in the magnetic field produced by the current. After using (52) to obtain several other general expressions for inductance, we shall show that it is equivalent to (49). We first express the potential energy W_H in terms of the magnetic fields,

$$L = \frac{\int_{\text{vol}} \mathbf{B} \cdot \mathbf{H}\, dv}{I^2} \tag{53}$$

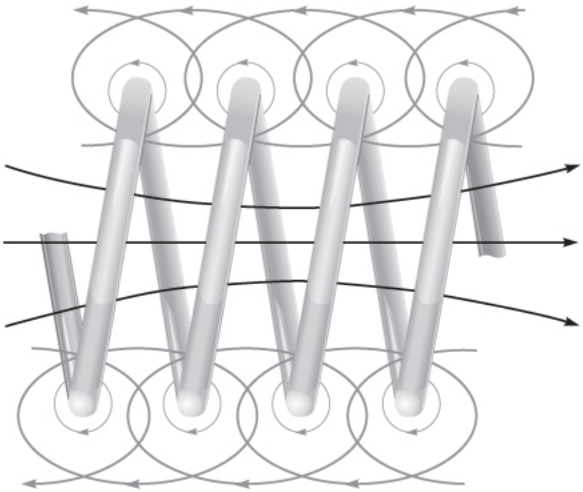


FIGURE 9.14
A portion of a coil showing partial flux linkages. The total flux linkages are obtained by adding the fluxes linking each turn.

and then replace $\mathbf{B}$ by $\nabla \times \mathbf{A}$,

$$L = \frac{1}{I^2} \int_{\text{vol}} \mathbf{H} \cdot (\nabla \times \mathbf{A}) dv$$

The vector identity

$$\nabla \cdot (\mathbf{A} \times \mathbf{H}) \equiv \mathbf{H} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{H}) \tag{54}$$

may be proved by expansion in cartesian coordinates. The inductance is then

$$L = \frac{1}{I^2} \left[\int_{\text{vol}} \nabla \cdot (\mathbf{A} \times \mathbf{H}) dv + \int_{\text{vol}} \mathbf{A} \cdot (\nabla \times \mathbf{H})\, dv \right] \tag{55}$$

After applying the divergence theorem to the first integral and letting $\nabla \times \mathbf{H} = \mathbf{J}$ in the second integral, we have

$$L = \frac{1}{I^2} \left[\oint_S (\mathbf{A} \times \mathbf{H}) \cdot d\mathbf{S} + \int_{\text{vol}} \mathbf{A} \cdot \mathbf{J}\, dv \right]$$

The surface integral is zero, since the surface encloses the volume containing all the magnetic energy, and this requires that $\mathbf{A}$ and $\mathbf{H}$ be zero on the bounding surface. The inductance may therefore be written as

$$L = \frac{1}{I^2} \int_{\text{vol}} \mathbf{A} \cdot \mathbf{J}\, dv \tag{56}$$

Equation (56) expresses the inductance in terms of an integral of the values of $\mathbf{A}$ and $\mathbf{J}$ at every point. Since current density exists only within the conductor, the integrand is zero at all points *outside* the conductor and the vector magnetic potential need not be determined there. The vector potential is that which arises from the current $\mathbf{J}$, and any other current source contributing a vector potential field in the region of the original current density is to be ignored for the present. Later we shall see that this leads to a *mutual inductance*.

The vector magnetic potential $\mathbf{A}$ due to $\mathbf{J}$ is given by Eq. (51), Chap. 8,

$$\mathbf{A} = \int_{\text{vol}} \frac{\mu \mathbf{J}}{4\pi R} dv$$

and the inductance may therefore be expressed more basically as a rather formidable double volume integral,

$$L = \frac{1}{I^2} \int_{\text{vol}} \left(\int_{\text{vol}} \frac{\mu \mathbf{J}}{4\pi R} dv \right) \cdot \mathbf{J}\, dv \tag{57}$$

A slightly simpler integral expression is obtained by restricting our attention to current filaments of small cross section for which $\mathbf{J}\, dv$ may be replaced by $I\, d\mathbf{L}$ and the volume integral by a closed line integral along the axis of the filament,

$$
\begin{aligned}
L &= \frac{1}{I^2} \oint \left(\oint \frac{\mu I \, d\mathbf{L}}{4\pi R} \right) \cdot I \, d\mathbf{L} \\
&= \frac{\mu}{4\pi} \oint \left(\oint \frac{d\mathbf{L}}{R} \right) \cdot d\mathbf{L}
\end{aligned} \tag{58}
$$

Our only present interest in Eqs. (57) and (58) lies in their implication that the inductance is a function of the distribution of the current in space or the geometry of the conductor configuration.

To obtain our original definition of inductance (49) let us hypothesize a uniform current distribution in a filamentary conductor of small cross section so that $\mathbf{J}\,dv$ in (56) becomes $I\,d\mathbf{L}$,

$$
L = \frac{1}{I} \oint \mathbf{A} \cdot d\mathbf{L} \tag{59}
$$

For a small cross section, $d\mathbf{L}$ may be taken along the center of the filament. We now apply Stokes' theorem and obtain

$$
L = \frac{1}{I} \int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S}
$$

or

$$
L = \frac{1}{I} \int_S \mathbf{B} \cdot d\mathbf{S}
$$

or

$$
L = \frac{\Phi}{I} \tag{60}
$$

Retracing the steps by which (60) is obtained, we should see that the flux Φ is that portion of the total flux which passes through any and every open surface whose perimeter is the filamentary current path.

If we now let the filament make N identical turns about the total flux, an idealization which may be closely realized in some types of inductors, the closed line integral must consist of N laps about this common path and (60) becomes

$$
L = \frac{N\Phi}{I} \tag{61}
$$

The flux Φ is now the flux crossing any surface whose perimeter is the path occupied by any *one* of the N turns. The inductance of an N-turn coil may still be obtained from (60), however, if we realize that the flux is that which crosses the complicated surface[4] whose perimeter consists of all N turns.

Use of any of the inductance expressions for a true filamentary conductor (having zero radius) leads to an infinite value of inductance, regardless of the configuration of the filament. Near the conductor Ampère's circuital law shows

[4] Somewhat like a spiral ramp.

that the magnetic field intensity varies inversely with the distance from the conductor, and a simple integration soon shows that an infinite amount of energy and an infinite amount of flux are contained within any finite cylinder about the filament. This difficulty is eliminated by specifying a small but finite filamentary radius.

The interior of any conductor also contains magnetic flux, and this flux links a variable fraction of the total current, depending on its location. These flux linkages lead to an *internal inductance*, which must be combined with the external inductance to obtain the total inductance. The internal inductance of a long straight wire of circular cross section, radius a, and uniform current distribution is

$$\boxed{L_{a,\text{int}} = \frac{\mu}{8\pi} \quad \text{H/m}} \tag{62}$$

a result requested in Prob. 43 at the end of this chapter.

In Chap. 11 it will be seen that the current distribution in a conductor at high frequencies tends to be concentrated near the surface. The internal flux is reduced, and it is usually sufficient to consider only the external inductance. At lower frequencies, however, internal inductance may become an appreciable part of the total inductance.

We conclude by defining the *mutual inductance* between circuits 1 and 2, M_{12}, in terms of mutual flux linkages,

$$\boxed{M_{12} = \frac{N_2 \Phi_{12}}{I_1}} \tag{63}$$

where Φ_{12} signifies the flux produced by I_1 which links the path of the filamentary current I_2, and N_2 is the number of turns in circuit 2. The mutual inductance, therefore, depends on the magnetic interaction between two currents. With either current alone, the total energy stored in the magnetic field can be found in terms of a single inductance, or self-inductance; with both currents having nonzero values, the total energy is a function of the two self-inductances and the mutual inductance. In terms of a mutual energy, it can be shown that (63) is equivalent to

$$M_{12} = \frac{1}{I_1 I_2} \int_{\text{vol}} (\mathbf{B}_1 \cdot \mathbf{H}_2) dv \tag{64}$$

or

$$M_{12} = \frac{1}{I_1 I_2} \int_{\text{vol}} (\mu \mathbf{H}_1 \cdot \mathbf{H}_2) dv \tag{65}$$

where $\mathbf{B}_1$ is the field resulting from I_1 (with $I_2 = 0$) and $\mathbf{H}_2$ is the field due to I_2 (with $I_1 = 0$). Interchange of the subscripts does not change the right-hand side of (65), and therefore

$$\boxed{M_{12} = M_{21}} \tag{66}$$

Mutual inductance is also measured in henrys, and we rely on the context to allow us to differentiate it from magnetization, also represented by M.

Example 9.9

Calculate the self-inductances of and the mutual inductances between two coaxial solenoids of radius R_1 and R_2, $R_2 > R_1$, carrying currents I_1 and I_2 with n_1 and n_2 turns/m, respectively.

Solution. We first attack the mutual inductances. From Eq. (15), Chap. 8, we let $n_1 = N/d$, and obtain

$$\begin{aligned} \mathbf{H}_1 &= n_1 I_1 \mathbf{a}_z \quad (0 < \rho < R_1) \\ &= 0 \quad (\rho > R_1) \end{aligned}$$

and

$$\begin{aligned} \mathbf{H}_2 &= n_2 I_2 \mathbf{a}_z \quad (0 < \rho < R_2) \\ &= 0 \quad (\rho > R_2) \end{aligned}$$

Thus, for this uniform field

$$\Phi_{12} = \mu_0 n_1 I_1 \pi R_1^2$$

and

$$M_{12} = \mu_0 n_1 n_2 \pi R_1^2$$

Similarly,

$$\Phi_{21} = \mu_0 n_2 I_2 \pi R_1^2$$
$$M_{21} = \mu_0 n_1 n_2 \pi R_1^2 = M_{12}$$

If $n_1 = 50$ turns/cm, $n_2 = 80$ turns/cm, $R_1 = 2$ cm, and $R_2 = 3$ cm, then

$$M_{12} = M_{21} = 4\pi \times 10^{-7}(5000)(8000)\pi(0.02^2) = 63.2 \quad \text{mH/m}$$

The self-inductances are easily found. The flux produced in coil 1 by I_1 is

$$\Phi_{11} = \mu_0 n_1 I_1 \pi R_1^2$$

and thus

$$L_1 = \mu_0 n_1^2 S_1 d \quad \text{H}$$

The inductance per unit length is therefore

$$L_1 = \mu_0 n_1^2 S_1 \quad \text{H/m}$$

or

$$L_1 = 39.5\,\mathrm{mH/m}$$

Similarly,

$$L_2 = \mu_0 n_2^2 S_2 = 22.7 \quad \mathrm{mH/m}$$

We see, therefore, that there are many methods available for the calculation of self-inductance and mutual inductance. Unfortunately, even problems possessing a high degree of symmetry present very challenging integrals for evaluation, and only a few problems are available for us to try our skill on.

Inductance will be discussed in circuit terms in Chap. 13.

✔ **D9.12.** Calculate the self-inductance of: (*a*) 3.5 m of coaxial cable with $a = 0.8$ mm and $b = 4$ mm, filled with a material for which $\mu_R = 50$; (*b*) a toroidal coil of 500 turns, wound on a fiberglass form having a 2.5×2.5 cm square cross section and an inner radius of 2 cm; (*c*) a solenoid having 500 turns about a cylindrical core of 2-cm radius in which $\mu_R = 50$ for $0 < \rho < 0.5$ cm and $\mu_R = 1$ for $0.5 < \rho < 2$ cm; the length of the solenoid is 50 cm.

Ans. 56.3 μH; 1.014 mH; 3.21 mH

✔ **D9.13.** A solenoid is 50-cm long, 2 cm in diameter, and contains 1500 turns. The cylindrical core has a diameter of 2 cm and a relative permeability of 75. This coil is coaxial with a second solenoid, also 50 cm long, but with a 3-cm diameter and 1200 turns. Calculate: (*a*) L for the inner solenoid; (*b*) L for the outer solenoid; (*c*) M between the two solenoids.

Ans. 133.2 mH; 86.7 mH; 106.6 mH

SUGGESTED REFERENCES

1. Kraus, J. D. and D. A. Fleisch: (see Suggested References for Chap. 3). Examples of the calculation of inductance are given on pp. 99–108.
2. Matsch, L. W.: (see Suggested References for Chap. 5). Chapter 3 is devoted to magnetic circuits and ferromagnetic materials.
3. Paul, C. R., K. W. Whites, and S. Y. Nasar: (see Suggested References for Chap. 8). Magnetic circuits, including those with permanent magnets, are discussed on pp. 263–270.

PROBLEMS

9.1 A point charge, $Q = -0.3\,\mu$C and $m = 3 \times 10^{-16}$ kg, is moving through the field $\mathbf{E} = 30\mathbf{a}_z$ V/m. Use Eq. (1) and Newton's laws to develop the appropriate differential equations and solve them, subject to the initial conditions at $t = 0$: $\mathbf{v} = 3 \times 10^5 \mathbf{a}_x$ m/s at the origin. At $t = 3\,\mu$s, find: (*a*) the position $P(x, y, z)$ of the charge; (*b*) the velocity $\mathbf{v}$; (*c*) the kinetic energy of the charge.

9.2 A point charge, $Q = -0.3\ \mu\text{C}$ and $m = 3 \times 10^{-16}$ kg, is moving through the field $\mathbf{B} = 30\mathbf{a}_z$ mT. Make use of Eq. (2) and Newton's laws to develop the appropriate differential equations and solve them, subject to the initial condition at $t = 0$, $\mathbf{v} = 3 \times 10^5 \mathbf{a}_x$ m/s at the origin. Solve these equations (perhaps with the help of an example given in Sec. 7.5) to evaluate at $t = 3\ \mu$s: (*a*) the position $P(x, y, z)$ of the charge; (*b*) its velocity; (*c*) and its kinetic energy.

9.3 A point charge for which $Q = 2 \times 10^{-16}$ C and $m = 5 \times 10^{-26}$ kg is moving in the combined fields $\mathbf{E} = 100\mathbf{a}_x - 200\mathbf{a}_y + 300\mathbf{a}_z$ V/m and $\mathbf{B} = -3\mathbf{a}_x + 2\mathbf{a}_y - \mathbf{a}_z$ mT. If the charge velocity at $t = 0$ is $\mathbf{v}(0) = (2\mathbf{a}_x - 3\mathbf{a}_y - 4\mathbf{a}_z)10^5$ m/s: (*a*) give the unit vector showing the direction in which the charge is accelerating at $t = 0$; (*b*) find the kinetic energy of the charge at $t = 0$.

9.4 An electron ($q_e = -1.60219 \times 10^{-19}$ C, $m = 9.10956 \times 10^{-31}$ kg) is moving at a constant velocity, $\mathbf{v} = 4.5 \times 10^7 \mathbf{a}_y$ m/s, along the negative y axis. At the origin, it encounters the uniform magnetic field $\mathbf{B} = 2.5\mathbf{a}_z$ mT, and remains in it up to $y = 2.5$ cm. If we assume (with good accuracy) that the electron remains on the y axis while it is in the magnetic field, find its x-, y-, and z-coordinate values when $y = 50$ cm.

9.5 A rectangular loop of wire in free space joins points $A(1, 0, 1)$ to $B(3, 0, 1)$ to $C(3, 0, 4)$ to $D(1, 0, 4)$ to A. The wire carries a current of 6 mA, flowing in the $\mathbf{a}_z$ direction from B to C. A filamentary current of 15 A flows along the entire z axis in the $\mathbf{a}_z$ direction. (*a*) Find $\mathbf{F}$ on side BC. (*b*) Find $\mathbf{F}$ on side AB. (*c*) Find $\mathbf{F}_{\text{total}}$ on the loop.

9.6 The magnetic flux density in a region of free space is given as $\mathbf{B} = -3x\mathbf{a}_x + 5y\mathbf{a}_y - 2z\mathbf{a}_z$ T. Find the total force on the rectangular loop shown in Fig. 9.15 if it lies in the plane $z = 0$ and is bounded by $x = 1$, $x = 3$, $y = 2$, and $y = 5$, all dimensions in cm.

9.7 Uniform current sheets are located in free space as follows: $8\mathbf{a}_z$ A/m at $y = 0$, $-4\mathbf{a}_z$ A/m at $y = 1$, and $-4\mathbf{a}_z$ A/m at $y = -1$. Find the

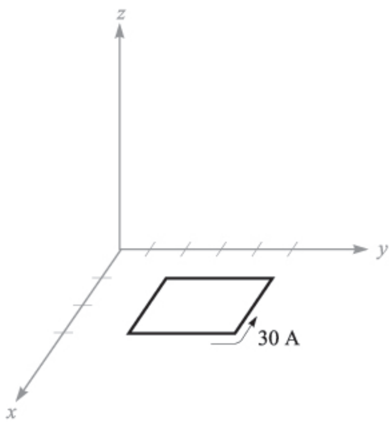


FIGURE 9.15
See Prob. 9.6.

vector force per meter length exerted on a current filament carrying 7 mA in the $\mathbf{a}_L$ direction if the filament is located at: (*a*) $x = 0$, $y = 0.5$, and $\mathbf{a}_L = \mathbf{a}_z$; (*b*) $y = 0.5$, $z = 0$, and $\mathbf{a}_L = \mathbf{a}_x$; (*c*) $x = 0$, $y = 1.5$, and $\mathbf{a}_L = \mathbf{a}_z$.

9.8 Filamentary currents of $-25\mathbf{a}_z$ and $25\mathbf{a}_z$ A are located in the $x = 0$ plane in free space at $y = -1$ and $y = 1$ m, respectively. A third filamentary current of $10^{-3}\mathbf{a}_z$ A is located at $x = k$, $y = 0$. Find the vector force on a 1-m length of the 1-mA filament and plot $|\mathbf{F}|$ versus k.

9.9 A current of $-100\mathbf{a}_z$ A/m flows on the conducting cylinder $\rho = 5$ mm and $+500\mathbf{a}_z$ A/m is present on the conducting cylinder $\rho = 1$ mm. Find the magnitude of the total force per meter length acting to split the outer cylinder apart along its length.

9.10 Two infinitely long parallel filaments each carry 50 A in the $\mathbf{a}_z$ direction. If the filaments lie in the plane $y = 0$ and $x = 5$ mm, find the vector force per meter length on the filament passing through the origin.

9.11 (*a*) Use Eq. (14), Sec. 9.3, to show that the force of attraction per unit length between two filamentary conductors in free space with currents $I_1\mathbf{a}_z$ at $x = 0$, $y = d/2$, and $I_2\mathbf{a}_z$ at $x = 0$, $y = -d/2$, is $\mu_0 I_1 I_2/(2\pi d)$. (*b*) Show how a simpler method can be used to check your result.

9.12 A conducting current strip carrying $\mathbf{K} = 12\mathbf{a}_z$ A/m lies in the $x = 0$ plane between $y = 0.5$ and $y = 1.5$ m. There is also a current filament of $I = 5$ A in the $\mathbf{a}_z$ direction on the z axis. Find the force exerted on the: (*a*) filament by the current strip; (*b*) strip by the filament.

9.13 A current of 6 A flows from $M(2, 0, 5)$ to $N(5, 0, 5)$ in a straight solid conductor in free space. An infinite current filament lies along the z axis and carries 50 A in the $\mathbf{a}_z$ direction. Compute the vector torque on the wire segment using an origin at: (*a*) (0,0,5); (*b*) (0,0,0); (*c*) (3,0,0).

9.14 The rectangular loop of Prob. 6 is now subjected to the $\mathbf{B}$ field produced by two current sheets, $\mathbf{K}_1 = 400\mathbf{a}_y$ A/m at $z = 2$, and $\mathbf{K}_2 = 300\mathbf{a}_z$ A/m at $y = 0$, in free space. Find the vector torque on the loop, referred to an origin: (*a*) at (0,0,0); (*b*) at the center of the loop.

9.15 A solid conducting filament extends from $x = -b$ to $x = b$ along the line $y = 2$, $z = 0$. This filament carries a current of 3 A in the $\mathbf{a}_x$ direction. An infinite filament on the z axis carries 5 A in the $\mathbf{a}_z$ direction. Obtain an expression for the torque exerted on the finite conductor about an origin located at (0,2,0).

9.16 Assume that an electron is describing a circular orbit of radius a about a positively charged nucleus. (*a*) By selecting an appropriate current and area, show that the equivalent orbital dipole moment is $ea^2\omega/2$, where ω is the electron's angular velocity. (*b*) Show that the torque produced by a magnetic field parallel to the plane of the orbit is $ea^2\omega B/2$. (*c*) By equating the Coulomb and centrifugal forces, show that ω is $(4\pi\epsilon_0 m_e a^3/e^2)^{-1/2}$, where m_e is the electron mass. (*d*) Find values for the angular velocity, torque, and the orbital magnetic moment for a hydrogen atom, where a is about 6×10^{-11} m; let $B = 0.5$ T.

9.17 The hydrogen atom described in Prob. 16 is now subjected to a magnetic field having the same direction as that of the atom. Show that the forces caused by B result in a decrease of the angular velocity by $eB/(2m_e)$ and a decrease in the orbital moment by $e^2a^2B/(4m_e)$. What are these decreases for the hydrogen atom in parts per million for an external magnetic flux density of 0.5 T?

9.18 Calculate the vector torque on the square loop shown in Fig. 9.16 about an origin at A in the field $\mathbf{B}$, given: (*a*) $A(0, 0, 0)$ and $\mathbf{B} = 100\mathbf{a}_y$ mT; (*b*) $A(0, 0, 0)$ and $\mathbf{B} = 200\mathbf{a}_x + 100\mathbf{a}_y$ mT; (*c*) $A(1, 2, 3)$ and $\mathbf{B} = 200\mathbf{a}_x + 100\mathbf{a}_y - 300\mathbf{a}_z$ mT; (*d*) $A(1, 2, 3)$ and $\mathbf{B} = 200\mathbf{a}_x + 100\mathbf{a}_y - 300\mathbf{a}_z$ mT for $x \geq 2$ and $\mathbf{B} = 0$ elsewhere.

9.19 Given a material for which $\chi_m = 3.1$ and within which $\mathbf{B} = 0.4y\mathbf{a}_z$ T, find: (*a*) $\mathbf{H}$; (*b*) μ; (*c*) μ_R; (*d*) $\mathbf{M}$; (*e*) $\mathbf{J}$; (*f*) $\mathbf{J}_b$; (*g*) $\mathbf{J}_T$.

9.20 Find $\mathbf{H}$ in a material where (*a*) $\mu_R = 4.2$, there are 2.7×10^{29} atoms/m^3, and each atom has a dipole moment of $2.6 \times 10^{-30}\mathbf{a}_y$ A $\cdot$ m^2; (*b*) $\mathbf{M} = 270\mathbf{a}_z$ A/m and $\mu = 2\,\mu$H/m; (*c*) $\chi_m = 0.7$ and $\mathbf{B} = 2\mathbf{a}_z$ T. (*d*) Find $\mathbf{M}$ in a material where bound surface current densities of $12\mathbf{a}_z$ A/m and $-9\mathbf{a}_z$ A/m exist at $\rho = 0.3$ m and 0.4 m, respectively.

9.21 Find the magnitude of the magnetization in a material for which: (*a*) the magnetic flux density is 0.02 Wb/m^2; (*b*) the magnetic field intensity is 1200 A/m and the relative permeability is 1.005; (*c*) there are 7.2×10^{28} atoms per cubic meter, each having a dipole moment of 4×10^{-30} A$\cdot$m^2 in the same direction, and the magnetic susceptibility is 0.0003.

9.22 Three current sheets are located as follows: $160\mathbf{a}_z$ A/m at $x = 1$ cm, $-40\mathbf{a}_z$ A/m at $x = 5$ cm, and $50\mathbf{a}_z$ A/m at $x = 8$ cm. Let $\mu = \mu_0$ for $x < 1$ cm and $x > 8$ cm; for $1 < x < 5$ cm, $\mu = 3\mu_0$, and for $5 < x < 8$ cm, $\mu = 2\mu_0$. Find $\mathbf{B}$ everywhere.

9.23 Calculate values for H_ϕ, B_ϕ, and M_ϕ at $\rho = c$ for a coaxial cable with $a = 2.5$ mm and $b = 6$ mm if it carries a current $I = 12$ A in the center conductor, and $\mu = 3\,\mu$H/m for 2.5 mm $< \rho <$ 3.5 mm, $\mu = 5\,\mu$H/m for

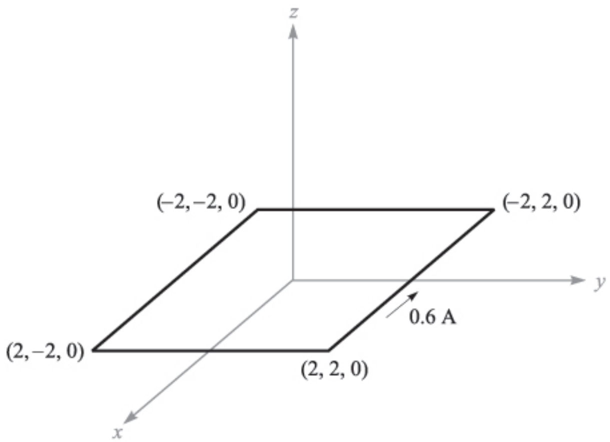


FIGURE 9.16
See Prob. 9.18.

$3.5\,\text{mm} < \rho < 4.5\,\text{mm}$, and $\mu = 10\,\mu\text{H/m}$ for $4.5\,\text{mm} < \rho < 6\,\text{mm}$. Use $c =$: (*a*) 3 mm; (*b*) 4 mm; (*c*) 5 mm.

9.24 A coaxial transmission line has $a = 5\,\text{mm}$ and $b = 20\,\text{mm}$. Let its center lie on the z axis and let a dc current I flow in the $\mathbf{a}_z$ direction in the center conductor. The volume between the conductors contains a magnetic material for which $\mu_R = 2.5$, as well as air. Find **H**, **B**, and **M** everywhere between conductors if $H_\phi = \frac{600}{\pi}$ A/m at $\rho = 10\,\text{mm}$, $\phi = \frac{\pi}{2}$, and the magnetic material is located where: (*a*) $a < \rho < 3a$; (*b*) $0 < \phi < \pi$.

9.25 A conducting filament at $z = 0$ carries 12 A in the $\mathbf{a}_z$ direction. Let $\mu_R = 1$ for $\rho < 1\,\text{cm}$, $\mu_R = 6$ for $1 < \rho < 2\,\text{cm}$, and $\mu_R = 1$ for $\rho > 2\,\text{cm}$. Find: (*a*) **H** everywhere; (*b*) **B** everywhere.

9.26 Point $P(2, 3, 1)$ lies on the planar boundary separating region 1 from region 2. The unit vector $\mathbf{a}_{N12} = 0.6\mathbf{a}_x + 0.48\mathbf{a}_y + 0.64\mathbf{a}_z$ is directed from region 1 to region 2. Let $\mu_{R1} = 2$, $\mu_{R2} = 8$, and $\mathbf{H}_1 = 100\mathbf{a}_x - 300\mathbf{a}_y + 200\mathbf{a}_z$ A/m. Find $\mathbf{H}_2$.

9.27 Let $\mu_{R1} = 2$ in region 1, defined by $2x + 3y - 4z > 1$, while $\mu_{R2} = 5$ in region 2 where $2x + 3y - 4z < 1$. In region 1, $\mathbf{H}_1 = 50\mathbf{a}_x - 30\mathbf{a}_y + 20\mathbf{a}_z$ A/m. Find: (*a*) $\mathbf{H}_{N1}$; (*b*) $\mathbf{H}_{t1}$; (*c*) $\mathbf{H}_{t2}$; (*d*) $\mathbf{H}_{N2}$; (*e*) θ_1, the angle between $\mathbf{H}_1$ and $\mathbf{a}_{N21}$; (*f*) θ_2, the angle between $\mathbf{H}_2$ and $\mathbf{a}_{N21}$.

9.28 For values of B below the knee on the magnetization curve for silicon steel approximate the curve by a straight line with $\mu = 5\,\text{mH/m}$. The core shown in Fig. 9.17 has areas of 1.6 cm^2 and lengths of 10 cm in each outer leg, and an area of 2.5 cm^2 and a length of 3 cm in the central leg. A coil of 1200 turns carrying 12 mA is placed around the central leg. Find B in the: (*a*) center leg; (*b*) center leg, if a 0.3-mm air gap is present in the center leg.

9.29 In Prob. 28, the linear approximation suggested in the statement of the problem leads to flux density of 0.666 T in the central leg. Using this value of B and the magnetization curve for silicon steel, what current is required in the 1200-turn coil?

9.30 A toroidal core has a circular cross section of 4 cm^2 area. The mean radius of the toroid is 6 cm. The core is composed of two semicircular segments, one of silicon steel and the other of a linear material with

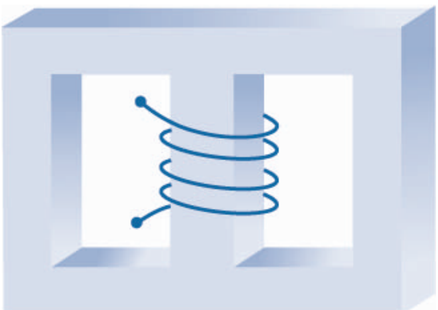


FIGURE 9.17
See Prob. 9.28.

$\mu_R = 200$. There is a 0.4-mm air gap at each of the two joints, and the core is wrapped by a 4000-turn coil carrying a dc current I_1. (*a*) Find I_1 if the flux density in the core is 1.2 T. (*b*) Find the flux density in the core if $I_1 = -.3$ A.

9.31 A toroid is constructed of a magnetic material having a cross-sectional area of 2.5 cm^2 and an effective length of 8 cm. There is also a short air gap of 0.25-mm length and an effective area of 2.8 cm^2. An mmf of 200 A·t is applied to the magnetic circuit. Calculate the total flux in the toroid if the magnetic material: (*a*) is assumed to have infinite permeability; (*b*) is assumed to be linear with $\mu_R = 1000$; (*c*) is silicon steel.

9.32 Determine the total energy stored in a spherical region 1 cm in radius, centered at the origin in free space, in the uniform field: (*a*) $\mathbf{H}_1 = -600\mathbf{a}_y$ A/m; (*b*) $600\mathbf{a}_x + 1200\mathbf{a}_y$ A/m; (*c*) $\mathbf{H}_3 = -600\mathbf{a}_x + 1200\mathbf{a}_y$ A/m; (*d*) $\mathbf{H}_4 = \mathbf{H}_2 + \mathbf{H}_3$; (*e*) $1000\mathbf{a}_x$ A/m + $0.001\mathbf{a}_x$ T.

9.33 A toroidal core has a square cross section, $2.5\,\text{cm} < \rho < 3.5\,\text{cm}$, $-0.5\,\text{cm} < z < 0.5\,\text{cm}$. The upper half of the toroid, $0 < z < 0.5\,\text{cm}$, is constructed of a linear material for which $\mu_R = 10$, while the lower half, $-0.5\,\text{cm} < z < 0$, has $\mu_R = 20$. An mmf of 150 A·t establishes a flux in the $\mathbf{a}_\phi$ direction. For $z > 0$, find: (*a*) $H_\phi(\rho)$; (*b*) $B_\phi(\rho)$; (*c*) $\Phi_{z>0}$. (*d*) Repeat for $z < 0$. (*e*) Find Φ_{total}.

9.34 Three planar current sheets are located in free space as follows: $-100\mathbf{a}_x$ A/m at $z = -1$, $200\mathbf{a}_x$ A/m at $z = 0$, and $-100\mathbf{a}_x$ A/m at $z = 1$. Let $w_H = \frac{1}{2}\mathbf{B}\cdot\mathbf{H}\,\text{J/m}^3$ and find w_H for all z.

9.35 The cones $\theta = 21^\circ$ and $\theta = 159^\circ$ are conducting surfaces and carry total currents of 40 A, as shown in Fig. 9.18. The currents return on a spherical conducting surface of 0.25-m radius. (*a*) Find $\mathbf{H}$ in the region $0 < r < 0.25$, $21^\circ < \theta < 159^\circ$, $0 < \phi < 2\pi$. (*b*) How much energy is stored in this region?

9.36 A filament carrying a current I in the $\mathbf{a}_z$ direction lies on the z axis, and cylindrical current sheets of $5\mathbf{a}_z$ A/m and $-2\mathbf{a}_z$ A/m are located at $\rho = 3$

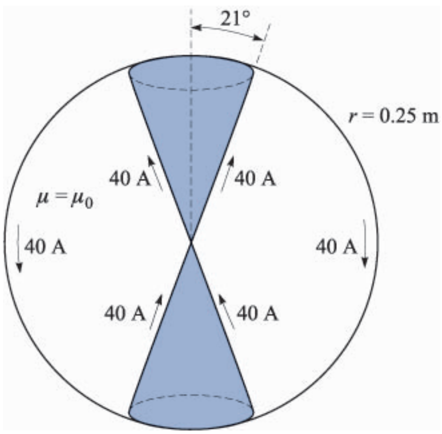


FIGURE 9.18
See Prob. 9.35.

and $\rho = 10$, respectively. (*a*) Find I if $\mathbf{H} = 0$ for $\rho > 10$. (*b*) Using this value of I, calculate $\mathbf{H}$ for all ρ, $3 < \rho < 10$. (*c*) Calculate and plot W_H versus ρ_0, where W_H is the total energy stored within the volume $0 < z < 1$, $0 < \phi < 2\pi$, $3 < \rho < \rho_0$.

9.37 Find the inductance of the cone-sphere configuration described in Prob. 35 and Fig. 9.18. The inductance is that offered at the origin between the vertices of the cone.

9.38 A toroidal core has a rectangular cross section defined by the surfaces $\rho = 2\,\text{cm}$, $\rho = 3\,\text{cm}$, $z = 4\,\text{cm}$, and $z = 4.5\,\text{cm}$. The core material has a relative permeability of 80. If the core is wound with a coil containing 8000 turns of wire, find its inductance.

9.39 Conducting planes in air at $z = 0$ and $z = d$ carry surface currents of $\pm K_0 \mathbf{a}_x$ A/m. (*a*) Find the energy stored in the magnetic field per unit length $(0 < x < 1)$ in a width $w\,(0 < y < w)$. (*b*) Calculate the inductance per unit length of this transmission line from $W_H = \frac{1}{2}LI^2$, where I is the total current in a width w in either conductor. (*c*) Calculate the total flux passing through the rectangle $0 < x < 1$, $0 < z < d$, in the plane $y = 0$, and from this result again find the inductance per unit length.

9.40 A coaxial cable has conductor dimensions of 1 and 5 mm. The region between conductors is air for $0 < \phi < \dfrac{\pi}{2}$ and $\pi < \phi < \dfrac{3\pi}{2}$, and a nonconducting material having $\mu_R = 8$ for $\dfrac{\pi}{2} < \phi < \pi$ and $\dfrac{3\pi}{2} < \phi < 2\pi$. Find the inductance per meter length.

9.41 A rectangular coil is composed of 150 turns of a filamentary conductor. Find the mutual inductance in free space between this coil and an infinite straight filament on the z axis if the four corners of the coil are located at: (*a*) (0,1,0), (0,3,0), (0,3,1), and (0,1,1); (*b*) (1,1,0), (1,3,0), (1,3,1), and (1,1,1).

9.42 Find the mutual inductance of this conductor system in free space: (*a*) the solenoid of Fig. 8.11*b* and a square filamentary loop of side length b coaxially centered inside the solenoid, if $a > b/\sqrt{2}$; (*b*) a cylindrical conducting shell of a radius a, axis on the z axis, and a filament at $x = 0$, $y = d$.

9.43 (*a*) Use energy relationships to show that the internal inductance of a nonmagnetic cylindrical wire of radius a carrying a uniformly distributed current I is $\mu_0/(8\pi)$ H/m. (*b*) Find the internal inductance if the portion of the conductor for which $\rho < c < a$ is removed.

CHAPTER 10

TIME-VARYING FIELDS AND MAXWELL'S EQUATIONS

The basic relationships of the electrostatic and the steady magnetic field were obtained in the previous nine chapters, and we are now ready to discuss time-varying fields. The discussion will be short, for vector analysis and vector calculus should now be more familiar tools; some of the relationships are unchanged, and most of the relationships are changed only slightly.

Two new concepts will be introduced: the electric field produced by a changing magnetic field and the magnetic field produced by a changing electric field. The first of these concepts resulted from experimental research by Michael Faraday, and the second from the theoretical efforts of James Clerk Maxwell.

Maxwell actually was inspired by Faraday's experimental work and by the mental picture provided through the "lines of force" that Faraday introduced in developing his theory of electricity and magnetism. He was 40 years younger than Faraday, but they knew each other during the 5 years Maxwell spent in London as a young professor, a few years after Faraday had retired. Maxwell's theory was developed subsequent to his holding this university position, while he was working alone at his home in Scotland. It occupied him for 5 years between the ages of 35 and 40.

The four basic equations of electromagnetic theory presented in this chapter bear his name.

10.1 FARADAY'S LAW

After Oersted[1] demonstrated in 1820 that an electric current affected a compass needle, Faraday professed his belief that if a current could produce a magnetic field, then a magnetic field should be able to produce a current. The concept of the "field" was not available at that time, and Faraday's goal was to show that a current could be produced by "magnetism."

He worked on this problem intermittently over a period of ten years, until he was finally successful in 1831.[2] He wound two separate windings on an iron toroid and placed a galvanometer in one circuit and a battery in the other. Upon closing the battery circuit, he noted a momentary deflection of the galvanometer; a similar deflection in the opposite direction occurred when the battery was disconnected. This, of course, was the first experiment he made involving a *changing* magnetic field, and he followed it with a demonstration that either a *moving* magnetic field or a moving coil could also produce a galvanometer deflection.

In terms of fields, we now say that a time-varying magnetic field produces an *electromotive force* (emf) which may establish a current in a suitable closed circuit. An electromotive force is merely a voltage that arises from conductors moving in a magnetic field or from changing magnetic fields, and we shall define it below. Faraday's law is customarily stated as

$$\boxed{\text{emf} = -\frac{d\Phi}{dt} \quad \text{V}} \tag{1}$$

Equation (1) implies a closed path, although not necessarily a closed conducting path; the closed path, for example, might include a capacitor, or it might be a purely imaginary line in space. The magnetic flux is that flux which passes through any and every surface whose perimeter is the closed path, and $d\Phi/dt$ is the time rate of change of this flux.

A nonzero value of $d\Phi/dt$ may result from any of the following situations:

1. A time-changing flux linking a stationary closed path
2. Relative motion between a steady flux and a closed path
3. A combination of the two

The minus sign is an indication that the emf is in such a direction as to produce a current whose flux, if added to the original flux, would reduce the magnitude of the emf. This statement that the induced voltage acts to produce an opposing flux is known as *Lenz's law*.[3]

---

[1] Hans Christian Oersted was Professor of Physics at the University of Copenhagen in Denmark.

[2] Joseph Henry produced similar results at Albany Academy in New York at about the same time.

[3] Henri Frederic Emile Lenz was born in Germany but worked in Russia. He published his law in 1834.

If the closed path is that taken by an N-turn filamentary conductor, it is often sufficiently accurate to consider the turns as coincident and let

$$\boxed{\text{emf} = -N\frac{d\Phi}{dt}} \tag{2}$$

where Φ is now interpreted as the flux passing through any one of N coincident paths.

We need to define emf as used in (1) or (2). The emf is obviously a scalar, and (perhaps not so obviously) a dimensional check shows that it is measured in volts. We define the emf as

$$\boxed{\text{emf} = \oint \mathbf{E} \cdot d\mathbf{L}} \tag{3}$$

and note that it is the voltage about a specific *closed path*. If any part of the path is changed, generally the emf changes. The departure from static results is clearly shown by (3), for an electric field intensity resulting from a static charge distribution must lead to zero potential difference about a closed path. In electrostatics, the line integral leads to a potential difference; with time-varying fields, the result is an emf or a voltage.

Replacing Φ in (1) by the surface integral of **B**, we have

$$\boxed{\text{emf} = \oint \mathbf{E} \cdot d\mathbf{L} = -\frac{d}{dt}\int_S \mathbf{B} \cdot d\mathbf{S}} \tag{4}$$

where the fingers of our right hand indicate the direction of the closed path, and our thumb indicates the direction of $d\mathbf{S}$. A flux density **B** in the direction of $d\mathbf{S}$ and increasing with time thus produces an average value of **E** which is *opposite* to the positive direction about the closed path. The right-handed relationship between the surface integral and the closed line integral in (4) should always be kept in mind during flux integrations and emf determinations.

Let us divide our investigation into two parts by first finding the contribution to the total emf made by a changing field within a stationary path (transformer emf), and then we will consider a moving path within a constant (motional, or generator, emf).

We first consider a stationary path. The magnetic flux is the only time-varying quantity on the right side of (4), and a partial derivative may be taken under the integral sign,

$$\text{emf} = \oint \mathbf{E} \cdot d\mathbf{L} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \tag{5}$$

Before we apply this simple result to an example, let us obtain the point form of this integral equation. Applying Stokes' theorem to the closed line integral, we have

$$\int_S (\nabla \times \mathbf{E}) \cdot d\mathbf{S} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}$$

where the surface integrals may be taken over identical surfaces. The surfaces are perfectly general and may be chosen as differentials,

$$(\nabla \times \mathbf{E}) \cdot d\mathbf{S} = -\frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}$$

and

$$\boxed{\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}} \tag{6}$$

This is one of Maxwell's four equations as written in differential, or point, form, the form in which they are most generally used. Equation (5) is the integral form of this equation and is equivalent to Faraday's law as applied to a fixed path. If **B** is not a function of time, (5) and (6) evidently reduce to the electrostatic equations,

$$\oint \mathbf{E} \cdot d\mathbf{L} = 0 \quad \text{(electrostatics)}$$

and

$$\nabla \times \mathbf{E} = 0 \quad \text{(electrostatics)}$$

As an example of the interpretation of (5) and (6), let us assume a simple magnetic field which increases exponentially with time within the cylindrical region $\rho < b$,

$$\mathbf{B} = B_0 e^{kt} \mathbf{a}_z \tag{7}$$

where $B_0 =$ constant. Choosing the circular path $\rho = a$, $a < b$, in the $z = 0$ plane, along which E_ϕ must be constant by symmetry, we then have from (5)

$$\text{emf} = 2\pi a E_\phi = -kB_0 e^{kt} \pi a^2$$

The emf around this closed path is $-kB_0 e^{kt}\pi a^2$. It is proportional to a^2, because the magnetic flux density is uniform and the flux passing through the surface at any instant is proportional to the area.

If we now replace a by ρ, $\rho < b$, the electric field intensity at any point is

$$\mathbf{E} = -\tfrac{1}{2} k B_0 e^{kt} \rho \mathbf{a}_\phi \tag{8}$$

Let us now attempt to obtain the same answer from (6), which becomes

$$(\nabla \times \mathbf{E})_z = -kB_0 e^{kt} = \frac{1}{\rho}\frac{\partial(\rho E_\phi)}{\partial \rho}$$

Multiplying by ρ and integrating from 0 to ρ (treating t as a constant, since the derivative is a partial derivative),

$$-\tfrac{1}{2}kB_0 e^{kt}\rho^2 = \rho E_\phi$$

or

$$\mathbf{E} = -\tfrac{1}{2}kB_0 e^{kt}\rho\mathbf{a}_\phi$$

once again.

If B_0 is considered positive, a filamentary conductor of resistance R would have a current flowing in the negative $\mathbf{a}_\phi$ direction, and this current would establish a flux within the circular loop in the negative $\mathbf{a}_z$ direction. Since E_ϕ increases exponentially with time, the current and flux do also, and thus tend to reduce the time rate of increase of the applied flux and the resultant emf in accordance with Lenz's law.

Before leaving this example, it is well to point out that the given field **B** does not satisfy all of Maxwell's equations. Such fields are often assumed (*always* in ac-circuit problems) and cause no difficulty when they are interpreted properly. They occasionally cause surprise, however. This particular field is discussed further in Prob. 19 at the end of the chapter.

Now let us consider the case of a time-constant flux and a moving closed path. Before we derive any special results from Faraday's law (1), let us use the basic law to analyze the specific problem outlined in Fig. 10.1. The closed circuit consists of two parallel conductors which are connected at one end by a high-resistance voltmeter of negligible dimensions and at the other end by a sliding bar moving at a velocity **v**. The magnetic flux density **B** is constant (in space and time) and is normal to the plane containing the closed path.

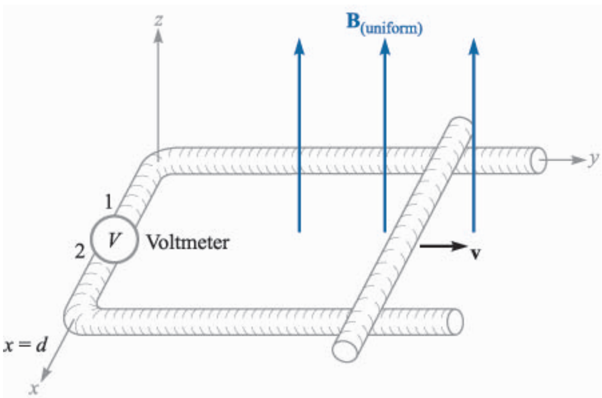


FIGURE 10.1
An example illustrating the application of Faraday's law to the case of a constant magnetic flux density **B** and a moving path. The shorting bar moves to the right with a velocity **v**, and the circuit is completed through the two rails and an extremely small high-resistance voltmeter. The voltmeter reading is $V_{12} = -Bvd$.

Let the position of the shorting bar be given by y; the flux passing through the surface within the closed path at any time t is then

$$\Phi = Byd$$

From (1), we obtain

$$\text{emf} = -\frac{d\Phi}{dt} = -B\frac{dy}{dt}d = -Bvd \tag{9}$$

The emf is defined as $\oint \mathbf{E} \cdot d\mathbf{L}$ and we have a conducting path; so we may actually determine $\mathbf{E}$ at every point along the closed path. We found in electrostatics that the tangential component of $\mathbf{E}$ is zero at the surface of a conductor, and we shall show in Sec. 10.4 that the tangential component is zero at the surface of a *perfect* conductor ($\sigma = \infty$) for all time-varying conditions. This is equivalent to saying that a perfect conductor is a "short circuit." The entire closed path in Figure 10.1 may be considered as a perfect conductor, with the exception of the voltmeter. The actual computation of $\oint \mathbf{E} \cdot d\mathbf{L}$ then must involve no contribution along the entire moving bar, both rails, and the voltmeter leads. Since we are integrating in a counterclockwise direction (keeping the interior of the positive side of the surface on our left as usual), the contribution $E\,\Delta L$ across the voltmeter must be $-Bvd$, showing that the electric field intensity in the instrument is directed from terminal 2 to terminal 1. For an up-scale reading, the positive terminal of the voltmeter should therefore be terminal 2.

The direction of the resultant small current flow may be confirmed by noting that the enclosed flux is reduced by a clockwise current in accordance with Lenz's law. The voltmeter terminal 2 is again seen to be the positive terminal.

Let us now consider this example using the concept of *motional emf*. The force on a charge Q moving at a velocity $\mathbf{v}$ in a magnetic field $\mathbf{B}$ is

$$\boxed{\mathbf{F} = Q\mathbf{v} \times \mathbf{B}}$$

or

$$\frac{\mathbf{F}}{Q} = \mathbf{v} \times \mathbf{B} \tag{10}$$

The sliding conducting bar is composed of positive and negative charges, and each experiences this force. The force per unit charge, as given by (10), is called the *motional* electric field intensity $\mathbf{E}_m$,

$$\boxed{\mathbf{E}_m = \mathbf{v} \times \mathbf{B}} \tag{11}$$

If the moving conductor were lifted off the rails, this electric field intensity would force electrons to one end of the bar (the far end) until the *static field* due to these

charges just balanced the field induced by the motion of the bar. The resultant tangential electric field intensity would then be zero along the length of the bar.

The motional emf produced by the moving conductor is then

$$\text{emf} = \oint \mathbf{E}_m \cdot d\mathbf{L} = \oint (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{L} \tag{12}$$

where the last integral may have a nonzero value only along that portion of the path which is in motion, or along which $\mathbf{v}$ has some nonzero value. Evaluating the right side of (12), we obtain

$$\oint (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{L} = \int_d^0 vB\,dx = -Bvd$$

as before. This is the total emf, since $\mathbf{B}$ is not a function of time.

In the case of a conductor moving in a uniform constant magnetic field, we may therefore ascribe a motional electric field intensity $\mathbf{E}_m = \mathbf{v} \times \mathbf{B}$ to every portion of the moving conductor and evaluate the resultant emf by

$$\text{emf} = \oint \mathbf{E} \cdot d\mathbf{L} = \oint \mathbf{E}_m \cdot d\mathbf{L} = \oint (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{L} \tag{13}$$

If the magnetic flux density is also changing with time, then we must include both contributions, the transformer emf (5) and the motional emf (12),

$$\text{emf} = \oint \mathbf{E} \cdot d\mathbf{L} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} + \oint (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{L} \tag{14}$$

This expression is equivalent to the simple statement

$$\text{emf} = -\frac{d\Phi}{dt} \tag{1}$$

and either can be used to determine these induced voltages.

Although (1) appears simple, there are a few contrived examples in which its proper application is quite difficult. These usually involve sliding contacts or switches; they always involve the substitution of one part of a circuit by a new part.[4] As an example, consider the simple circuit of Fig. 10.2, containing several perfectly conducting wires, an ideal voltmeter, a uniform constant field $\mathbf{B}$, and a switch. When the switch is opened, there is obviously more flux enclosed in the voltmeter circuit; however, it continues to read zero. The change in flux has not been produced by either a time-changing $\mathbf{B}$ [first term of (14)] or a conductor moving through a magnetic field [second part of (14)]. Instead, a new circuit has been substituted for the old. Thus it is necessary to use care in evaluating the change in flux linkages.

The separation of the emf into the two parts indicated by (14), one due to the time rate of change of $\mathbf{B}$ and the other to the motion of the circuit, is some-

[4] See Bewley, in Suggested References at the end of the chapter, particularly pp. 12–19.

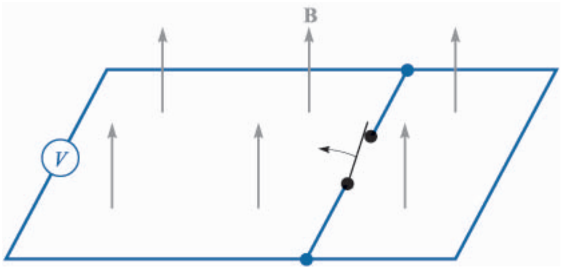


FIGURE 10.2
An apparent increase in flux linkages does not lead to an induced voltage when one part of a circuit is simply substituted for another by opening the switch. No indication will be observed on the voltmeter.

what arbitrary in that it depends on the relative velocity of the *observer* and the system. A field that is changing with both time and space may look constant to an observer moving with the field. This line of reasoning is developed more fully in applying the special theory of relativity to electromagnetic theory.[5]

✔ **D10.1.** Within a certain region, $\epsilon = 10^{-11}$ F/m and $\mu = 10^{-5}$ H/m. If $B_x = 2 \times 10^{-4} \cos 10^5 t \sin 10^{-3} y$ T: (*a*) use $\nabla \times \mathbf{H} = \epsilon \frac{\partial \mathbf{E}}{\partial t}$ to find **E**; (*b*) find the total magnetic flux passing through the surface $x = 0$, $0 < y < 40$ m, $0 < z < 2$ m, at $t = 1\ \mu$s; (*c*) find the value of the closed line integral of **E** around the perimeter of the given surface.

Ans. $-20\,000 \sin 10^5 t \cos 10^{-3} y \mathbf{a}_z$ V/m; 31.4 mWb; -315 V

✔ **D10.2.** With reference to the sliding bar shown in Figure 10.1, let $d = 7$ cm, $\mathbf{B} = 0.3\mathbf{a}_z$ T, and $\mathbf{v} = 0.1\mathbf{a}_y e^{20y}$ m/s. Let $y = 0$ at $t = 0$. Find: (*a*) $v(t = 0)$; (*b*) $y(t = 0.1)$; (*c*) $v(t = 0.1)$; (*d*) V_{12} at $t = 0.1$.

Ans. 0.1 m/s; 1.116 cm; 0.1250 m/s; -0.002625 V

10.2 DISPLACEMENT CURRENT

Faraday's experimental law has been used to obtain one of Maxwell's equations in differential form,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{15}$$

which shows us that a time-changing magnetic field produces an electric field. Remembering the definition of curl, we see that this electric field has the special property of circulation; its line integral about a general closed path is not zero. Now let us turn our attention to the time-changing electric field.

We should first look at the point form of Ampère's circuital law as it applies to steady magnetic fields,

[5] This is discussed in several of the references listed in the Suggested References at the end of the chapter. See Panofsky and Phillips, pp. 142–151; Owen, pp. 231–245; and Harman in several places.

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{16}$$

and show its inadequacy for time-varying conditions by taking the divergence of each side,

$$\nabla \cdot \nabla \times \mathbf{H} \equiv 0 = \nabla \cdot \mathbf{J}$$

The divergence of the curl is identically zero, so $\nabla \cdot \mathbf{J}$ is also zero. However, the equation of continuity,

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho_v}{\partial t}$$

then shows us that (16) can be true only if $\partial \rho_v / \partial t = 0$. This is an unrealistic limitation, and (16) must be amended before we can accept it for time-varying fields. Suppose we add an unknown term $\mathbf{G}$ to (16),

$$\nabla \times \mathbf{H} = \mathbf{J} + \mathbf{G}$$

Again taking the divergence, we have

$$0 = \nabla \cdot \mathbf{J} + \nabla \cdot \mathbf{G}$$

Thus

$$\nabla \cdot \mathbf{G} = \frac{\partial \rho_v}{\partial t}$$

Replacing ρ_v by $\nabla \cdot \mathbf{D}$,

$$\nabla \cdot \mathbf{G} = \frac{\partial}{\partial t}(\nabla \cdot \mathbf{D}) = \nabla \cdot \frac{\partial \mathbf{D}}{\partial t}$$

from which we obtain the simplest solution for $\mathbf{G}$,

$$\mathbf{G} = \frac{\partial \mathbf{D}}{\partial t}$$

Ampère's circuital law in point form therefore becomes

$$\boxed{\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}} \tag{17}$$

Equation (17) has not been derived. It is merely a form we have obtained which does not disagree with the continuity equation. It is also consistent with all our other results, and we accept it as we did each experimental law and the equations derived from it. We are building a theory, and we have every right to our equations *until they are proved wrong*. This has not yet been done.

We now have a second one of Maxwell's equations and shall investigate its significance. The additional term $\partial \mathbf{D}/\partial t$ has the dimensions of current density, amperes per square meter. Since it results from a time-varying electric flux density (or displacement density), Maxwell termed it a *displacement current density*. We sometimes denote it by $\mathbf{J}_d$:

$$\nabla \times \mathbf{H} = \mathbf{J} + \mathbf{J}_d$$

$$\mathbf{J}_d = \frac{\partial \mathbf{D}}{\partial t}$$

This is the third type of current density we have met. Conduction current density,

$$\mathbf{J} = \sigma \mathbf{E}$$

is the motion of charge (usually electrons) in a region of zero net charge density, and convection current density,

$$\mathbf{J} = \rho_v \mathbf{v}$$

is the motion of volume charge density. Both are represented by **J** in (17). Bound current density is, of course, included in **H**. In a nonconducting medium in which no volume charge density is present, $\mathbf{J} = 0$, and then

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \quad (\text{if } \mathbf{J} = 0) \tag{18}$$

Notice the symmetry between (18) and (15):

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{15}$$

Again the analogy between the intensity vectors **E** and **H** and the flux density vectors **D** and **B** is apparent. Too much faith cannot be placed in this analogy, however, for it fails when we investigate forces on particles. The force on a charge is related to **E** and to **B**, and some good arguments may be presented showing an analogy between **E** and **B** and between **D** and **H**. We shall omit them, however, and merely say that the concept of displacement current was probably suggested to Maxwell by the symmetry first mentioned above.[6]

The total displacement current crossing any given surface is expressed by the surface integral,

$$I_d = \int_S \mathbf{J}_d \cdot d\mathbf{S} = \int_S \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{S}$$

and we may obtain the time-varying version of Ampère's circuital law by integrating (17) over the surface S,

$$\int_S (\nabla \times \mathbf{H}) \cdot d\mathbf{S} = \int_S \mathbf{J} \cdot d\mathbf{S} + \int_S \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{S}$$

[6] The analogy that relates **B** to **D** and **H** to **E** is strongly advocated by Fano, Chu, and Adler (see Suggested References for Chap. 5) on pp. 159–160 and 179; the case for comparing **B** to **E** and **D** to **H** is presented in Halliday and Resnick (see Suggested References for this chapter) on pp. 665–668 and 832–836.

and applying Stokes' theorem,

$$\oint \mathbf{H} \cdot d\mathbf{L} = I + I_d = I + \int_S \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{S} \tag{19}$$

What is the nature of displacement current density? Let us study the simple circuit of Fig. 10.3, containing a filamentary loop and a parallel-plate capacitor. Within the loop a magnetic field varying sinusoidally with time is applied to produce an emf about the closed path (the filament plus the dashed portion between the capacitor plates) which we shall take as

$$\text{emf} = V_0 \cos \omega t$$

Using elementary circuit theory and assuming the loop has negligible resistance and inductance, we may obtain the current in the loop as

$$I = -\omega C V_0 \sin \omega t$$
$$= -\omega \frac{\epsilon S}{d} V_0 \sin \omega t$$

where the quantities ϵ, S, and d pertain to the capacitor. Let us apply Ampère's circuital law about the smaller closed circular path k and neglect displacement current for the moment:

$$\oint_k \mathbf{H} \cdot d\mathbf{L} = I_k$$

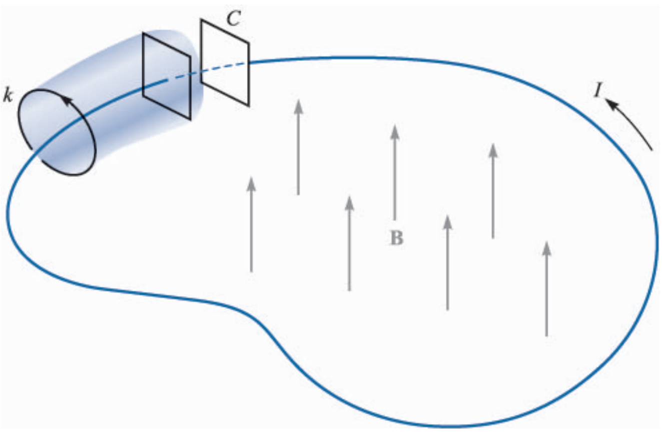


FIGURE 10.3
A filamentary conductor forms a loop connecting the two plates of a parallel-plate capacitor. A time-varying magnetic field inside the closed path produces an emf of $V_0 \cos \omega t$ around the closed path. The conduction current I is equal to the displacement current between the capacitor plates.

The path and the value of **H** along the path are both definite quantities (although difficult to determine), and $\oint_k \mathbf{H} \cdot d\mathbf{L}$ is a definite quantity. The current I_k is that current through every surface whose perimeter is the path k. If we choose a simple surface punctured by the filament, such as the plane circular surface defined by the circular path k, the current is evidently the conduction current. Suppose now we consider the closed path k as the mouth of a paper bag whose bottom passes between the capacitor plates. The bag is not pierced by the filament, and the conductor current is zero. Now we need to consider displacement current, for within the capacitor

$$D = \epsilon E = \epsilon\left(\frac{V_0}{d}\cos\omega t\right)$$

and therefore

$$I_d = \frac{\partial D}{\partial t} S = -\omega \frac{\epsilon S}{d} V_0 \sin\omega t$$

This is the same value as that of the conduction current in the filamentary loop. Therefore the applicaton of Ampère's circuital law including displacement current to the path k leads to a definite value for the line integral of **H**. This value must be equal to the total current crossing the chosen surface. For some surfaces the current is almost entirely conduction current, but for those surfaces passing between the capacitor plates, the conduction current is zero, and it is the displacement current which is now equal to the closed line integral of **H**.

Physically, we should note that a capacitor stores charge and that the electric field between the capacitor plates is much greater than the small leakage fields outside. We therefore introduce little error when we neglect displacement current on all those surfaces which do not pass between the plates.

Displacement current is associated with time-varying electric fields and therefore exists in all imperfect conductors carrying a time-varying conduction current. The last part of the drill problem below indicates the reason why this additional current was never discovered experimentally. This comparison is illustrated further in Sec. 11.3.

✔ **D10.3.** Find the amplitude of the displacement current density: (*a*) adjacent to an automobile antenna where the magnetic field intensity of an FM signal is $H_x = 0.15\cos[3.12(3\times10^8 t - y)]$ A/m; (*b*) in the air space at a point within a large power distribution transformer where $\mathbf{B} = 0.8\cos[1.257\times10^{-6}(3\times10^8 t - x)]\mathbf{a}_y$ T; (*c*) within a large oil-filled power capacitor where $\epsilon_R = 5$ and $\mathbf{E} = 0.9\cos[1.257\times10^{-6}(3\times10^8 t - z\sqrt{5})]\mathbf{a}_x$ MV/m; (*d*) in a metallic conductor at 60 Hz, if $\epsilon = \epsilon_0$, $\mu = \mu_0$, $\sigma = 5.8\times10^7$ S/m, and $\mathbf{J} = \sin(377t - 117.1z)\mathbf{a}_x$ MA/m^2.

Ans. 0.318 A/m^2; 0.800 A/m^2; 0.01502 A/m^2; 57.6 pA/m^2

10.3 MAXWELL'S EQUATIONS IN POINT FORM

We have already obtained two of Maxwell's equations for time-varying fields,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{20}$$

and

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{21}$$

The remaining two equations are unchanged from their non-time-varying form:

$$\nabla \cdot \mathbf{D} = \rho_v \tag{22}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{23}$$

Equation (22) essentially states that charge density is a source (or sink) of electric flux lines. Note that we can no longer say that *all* electric flux begins and terminates on charge, because the point form of Faraday's law (20) shows that **E**, and hence **D**, may have circulation if a changing magnetic field is present. Thus the lines of electric flux may form closed loops. However, the converse is still true, and every coulomb of charge must have one coulomb of electric flux diverging from it.

Equation (23) again acknowledges the fact that "magnetic charges," or poles, are not known to exist. Magnetic flux is always found in closed loops and never diverges from a point source.

These for equations form the basis of all electromagnetic theory. They are partial differential equations and relate the electric and magnetic fields to each other and to their sources, charge and current density. The auxiliary equations relating **D** and **E**.

$$\mathbf{D} = \epsilon \mathbf{E} \tag{24}$$

relating **B** and **H**,

$$\mathbf{B} = \mu \mathbf{H} \tag{25}$$

defining conduction current density,

$$\boxed{\mathbf{J} = \sigma \mathbf{E}} \tag{26}$$

and defining convection current density in terms of the volume charge density ρ_v,

$$\boxed{\mathbf{J} = \rho_v \mathbf{v}} \tag{27}$$

are also required to define and relate the quantities appearing in Maxwell's equations.

The potentials V and $\mathbf{A}$ have not been included above because they are not strictly necessary, although they are extremely useful. They will be discussed at the end of this chapter.

If we do not have "nice" materials to work with, then we should replace (24) and (25) by the relationships involving the polarization and magnetization fields,

$$\boxed{\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}} \tag{28}$$

$$\boxed{\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M})} \tag{29}$$

For linear materials we may relate $\mathbf{P}$ to $\mathbf{E}$

$$\mathbf{P} = \chi_e \epsilon_0 \mathbf{E} \tag{30}$$

and $\mathbf{M}$ to $\mathbf{H}$

$$\mathbf{M} = \chi_m \mathbf{H} \tag{31}$$

Finally, because of its fundamental importance we should include the Lorentz force equation, written in point form as the force per unit volume,

$$\boxed{\mathbf{f} = \rho_v (\mathbf{E} + \mathbf{v} \times \mathbf{B})} \tag{32}$$

The following chapters are devoted to the application of Maxwell's equations to several simple problems.

✔ **D10.4.** Let $\mu = 10^{-5}$ H/m, $\epsilon = 4 \times 10^{-9}$ F/m, $\sigma = 0$, and $\rho_v = 0$. Find k (including units) so that each of the following pairs of fields satisfies Maxwell's equations: (*a*) $\mathbf{D} = 6\mathbf{a}_x - 2y\mathbf{a}_y + 2z\mathbf{a}_z$ nC/m^2, $\mathbf{H} = kx\mathbf{a}_x + 10y\mathbf{a}_y - 25z\mathbf{a}_z$ A/m; (*b*) $\mathbf{E} = (20y - kt)\mathbf{a}_x$ V/m, $\mathbf{H} = (y + 2 \times 10^6 t)\mathbf{a}_z$ A/m.

Ans. 15 A/m^2; -2.5×10^8 V/(m·s)

10.4 MAXWELL'S EQUATIONS IN INTEGRAL FORM

The integral forms of Maxwell's equations are usually easier to recognize in terms of the experimental laws from which they have been obtained by a generalization process. Experiments must treat physical macroscopic quantities, and their results therefore are expressed in terms of integral relationships. A differential equation always represents a theory. Let us now collect the integral forms of Maxwell's equations of the previous section.

Integrating (20) over a surface and applying Stokes' theorem, we obtain Faraday's law,

$$\boxed{\oint \mathbf{E} \cdot d\mathbf{L} = -\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}} \qquad (33)$$

and the same process applied to (21) yields Ampère's circuital law,

$$\boxed{\oint \mathbf{H} \cdot d\mathbf{L} = I + \int_S \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{S}} \qquad (34)$$

Gauss's laws for the electric and magnetic fields are obtained by integrating (22) and (23) throughout a volume and using the divergence theorem:

$$\boxed{\oint_S \mathbf{D} \cdot d\mathbf{S} = \int_{\text{vol}} \rho_v dv} \qquad (35)$$

$$\boxed{\oint_S \mathbf{B} \cdot d\mathbf{S} = 0} \qquad (36)$$

These four integral equations enable us to find the boundary conditions on **B**, **D**, **H**, and **E** which are necessary to evaluate the constants obtained in solving Maxwell's equations in partial differential form. These boundary conditions are in general unchanged from their forms for static or steady fields, and the same methods may be used to obtain them. Between any two real physical media (where **K** must be zero on the boundary surface), (33) enables us to relate the tangential **E**-field components,

$$E_{t1} = E_{t2} \qquad (37)$$

and from (34),

$$H_{t1} = H_{t2} \qquad (38)$$

The surface integrals produce the boundary conditions on the normal components,

$$D_{N1} - D_{N2} = \rho_S \tag{39}$$

and

$$B_{N1} = B_{N2} \tag{40}$$

It is often desirable to idealize a physical problem by assuming a perfect conductor for which σ is infinite but **J** is finite. From Ohm's law, then, in a perfect conductor,

$$\mathbf{E} = 0$$

and it follows from the point form of Faraday's law that

$$\mathbf{H} = 0$$

for time-varying fields. The point form of Ampère's circuital law then shows that the finite value of **J** is

$$\mathbf{J} = 0$$

and current must be carried on the conductor surface as a surface current **K**. Thus, if region 2 is a perfect conductor, (37) to (40) become, respectively,

$$E_{t1} = 0 \tag{41}$$

$$H_{t1} = K \quad (\mathbf{H}_{t1} = \mathbf{K} \times \mathbf{a}_N) \tag{42}$$

$$D_{N1} = \rho_s \tag{43}$$

$$B_{N1} = 0 \tag{44}$$

where $\mathbf{a}_N$ is an outward normal at the conductor surface.

Note that surface charge density is considered a physical possibility for either dielectrics, perfect conductors, or imperfect conductors, but that surface *current* density is assumed only in conjunction with perfect conductors.

The boundary conditions stated above are a very necessary part of Maxwell's equations. All real physical problems have boundaries and require the solution of Maxwell's equations in two or more regions and the matching of these solutions at the boundaries. In the case of perfect conductors, the solution of the equations within the conductor is trivial (all time-varying fields are zero), but the application of the boundary conditions (41) to (44) may be very difficult.

Certain fundamental properties of wave propagation are evident when Maxwell's equations are solved for an *unbounded* region. This problem is treated in the following chapter. It represents the simplest application of Maxwell's equations, becuase it is the only problem which does not require the application of any boundary conditions.

✔ **D10.5.** The unit vector $0.64\mathbf{a}_x + 0.6\mathbf{a}_y - 0.48\mathbf{a}_z$ is directed from region 2 ($\epsilon_R = 2, \mu_R = 3, \sigma_2 = 0$) toward region 1 ($\epsilon_{R1} = 4, \mu_{R1} = 2, \sigma_1 = 0$). If $\mathbf{B}_1 = (\mathbf{a}_x - 2\mathbf{a}_y + 3\mathbf{a}_z)\sin 300t$ T at point P in region 1 adjacent to the boundary, find the amplitude at P of: (*a*) $\mathbf{B}_{N1}$; (*b*) $\mathbf{B}_{t1}$; (*c*) $\mathbf{B}_{N2}$; (*d*) $\mathbf{B}_2$.

Ans. 2.00 T; 3.26 T; 2.00 T; 5.15 T

✔ **D10.6.** The surface $y = 0$ is a perfectly conducting plane, while the region $y > 0$ has $\epsilon_R = 5$, $\mu_R = 3$, and $\sigma = 0$. Let $\mathbf{E} = 20\cos(2\times 10^8 t - 2.58z)\mathbf{a}_y$ V/m for $y > 0$, and find at $t = 6$ ns; (*a*) ρ_S at $P(2, 0, 0.3)$; (*b*) $\mathbf{H}$ at P; (*c*) $\mathbf{K}$ at P.

Ans. 40.3 nC/m^2; $-62.3\mathbf{a}_x$ mA/m; $-62.3\mathbf{a}_z$ mA/m

10.5 THE RETARDED POTENTIALS

The time-varying potentials, usually called *retarded* potentials for a reason which we shall see shortly, find their greatest application in radiation problems in which the distribution of the source is known approximately. We should remember that the scalar electric potential V may be expressed in terms of a static charge distribution,

$$V = \int_{\text{vol}} \frac{\rho_v dv}{4\pi\epsilon R} \quad \text{(static)} \tag{45}$$

and the vector magnetic potential may be found from a current distribution which is constant with time,

$$\mathbf{A} = \int_{\text{vol}} \frac{\mu \mathbf{J}\, dv}{4\pi R} \quad \text{(dc)} \tag{46}$$

The differential equations satisfied by V,

$$\nabla^2 V = -\frac{\rho_v}{\epsilon} \quad \text{(static)} \tag{47}$$

and $\mathbf{A}$,

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J} \quad \text{(dc)} \tag{48}$$

may be regarded as the point forms of the integral equations (45) and (46), respectively.

Having found V and $\mathbf{A}$, the fundamental fields are then simply obtained by using the gradient,

$$\mathbf{E} = -\nabla V \quad \text{(static)} \tag{49}$$

or the curl,

$$\mathbf{B} = \nabla \times \mathbf{A} \quad \text{(dc)} \tag{50}$$

We now wish to define suitable time-varying potentials which are consistent with the above expressions when only static charges and direct currents are involved.

Equation (50) apparently is still consistent with Maxwell's equations. These equations state that $\nabla \cdot \mathbf{B} = 0$, and the divergence of (50) leads to the divergence of the curl which is identically zero. Let us therefore tentatively accept (50) as satisfactory for time-varying fields and turn our attention to (49).

The inadequacy of (49) is obvious, because application of the curl operation to each side and recognition of the curl of the gradient as being identically zero confront us with $\nabla \times \mathbf{E} = 0$. However, the point form of Faraday's law states that $\nabla \times \mathbf{E}$ is not generally zero, so let us try to effect an improvement by adding an unknown term to (49),

$$\mathbf{E} = -\nabla V + \mathbf{N}$$

taking the curl,

$$\nabla \times \mathbf{E} = 0 + \nabla \times \mathbf{N}$$

using the point form of Faraday's law,

$$\nabla \times \mathbf{N} = -\frac{\partial \mathbf{B}}{\partial t}$$

and using (50), giving us

$$\nabla \times \mathbf{N} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{A})$$

or

$$\nabla \times \mathbf{N} = -\nabla \times \frac{\partial \mathbf{A}}{\partial t}$$

The simplest solution of this equation is

$$\mathbf{N} = -\frac{\partial \mathbf{A}}{\partial t}$$

and this leads to

$$\boxed{\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}} \tag{51}$$

We still must check (50) and (51) by substituting them into the remaining two of Maxwell's equations:

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

$$\nabla \cdot \mathbf{D} = \rho_v$$

Doing this, we obtain the more complicated expressions

$$\frac{1}{\mu}\nabla \times \nabla \times \mathbf{A} = \mathbf{J} + \epsilon\left(-\nabla\frac{\partial V}{\partial t} - \frac{\partial^2 \mathbf{A}}{\partial t^2}\right)$$

and

$$\epsilon\left(-\nabla \cdot \nabla V - \frac{\partial}{\partial t}\nabla \cdot \mathbf{A}\right) = \rho_v$$

or

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mu \mathbf{J} - \mu\epsilon\left(\nabla\frac{\partial V}{\partial t} + \frac{\partial^2 \mathbf{A}}{\partial t^2}\right) \tag{52}$$

and

$$\nabla^2 V + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) = -\frac{\rho_v}{\epsilon} \tag{53}$$

There is no apparent inconsistency in (52) and (53). Under static or dc conditions $\nabla \cdot \mathbf{A} = 0$, and (52) and (53) reduce to (48) and (47), respectively. We shall therefore assume that the time-varying potentials may be defined in such a way that **B** and **E** may be obtained from them through (50) and (51). These latter two equations do not serve, however, to define **A** and V *completely*. They represent necessary, but not sufficient, conditions. Our initial assumption was merely that $B = \nabla \times \mathbf{A}$, and a vector cannot be defined by giving its curl alone. Suppose, for example, that we have a very simple vector potential field in which A_y and A_z are zero. Expansion of (50) leads to

$$B_x = 0$$
$$B_y = \frac{\partial A_x}{\partial z}$$
$$B_z = -\frac{\partial A_x}{\partial y}$$

and we see that no information is available about the manner in which A_x varies with x. This information could be found if we also knew the value of the divergence of **A**, for in our example

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x}$$

Finally, we should note that our information about **A** is given only as partial derivatives and that a space-constant term might be added. In all physical problems in which the region of the solution extends to infinity, this constant term must be zero, for there can be no fields at infinity.

Generalizing from this simple example, we may say that a vector field is defined completely when both its curl and divergence are given and when its value is known at any one point (including infinity). We are therefore at liberty

to specify the divergence of **A**, and we do so with an eye on (52) and (53), seeking the simplest expressions. We define

$$\nabla \cdot \mathbf{A} = -\mu\epsilon \frac{\partial V}{\partial t} \tag{54}$$

and (52) and (53) become

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J} + \mu\epsilon \frac{\partial^2 \mathbf{A}}{\partial t^2} \tag{55}$$

and

$$\nabla^2 V = -\frac{\rho_v}{\epsilon} + \mu\epsilon \frac{\partial^2 V}{\partial t^2} \tag{56}$$

These equations are related to the wave equation, which will be discussed in the following chapter. They show considerable symmetry, and we should be highly pleased with our definitions of V and **A**,

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{50}$$

$$\nabla \cdot \mathbf{A} = -\mu\epsilon \frac{\partial V}{\partial t} \tag{54}$$

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \tag{51}$$

The integral equivalents of (45) and (46) for the time-varying potentials follow from the definitions (50), (51), and (54), but we shall merely present the final results and indicate their general nature. In the next chapter a study of the uniform plane wave will introduce the concept of *propagation*, in which any electromagnetic disturbance is found to travel at a velocity

$$v = \frac{1}{\sqrt{\mu\epsilon}}$$

through any homogeneous medium described by μ and ϵ. In the case of free space this velocity turns out to be velocity of light, approximately 3×10^8 m/s. It is logical, then, to suspect that the potential at any point is due not to the value of the charge density at some distant point at the same instant, but to its value at some previous time, because the effect propagates at a finite velocity. Thus (45) becomes

$$\boxed{V = \int_{\text{vol}} \frac{[\rho_v]}{4\pi\epsilon R} dv} \tag{57}$$

where $[\rho_v]$ indicates that every t appearing in the expression for ρ_v has been replaced by a *retarded* time,

$$t' = t - \frac{R}{v}$$

Thus, if the charge density throughout space were given by

$$\rho_v = e^{-r} \cos \omega t$$

then

$$[\rho_v] = e^{-r} \cos\left[\omega\left(t - \frac{R}{v}\right)\right]$$

where R is the distance between the differential element of charge being considered and the point at which the potential is to be determined.

The retarded vector magnetic potential is given by

$$\mathbf{A} = \int_{\text{vol}} \frac{\mu[\mathbf{J}]}{4\pi R} dv \tag{58}$$

The use of a retarded time has resulted in the time-varying potentials being given the name of retarded potentials. In Chap. 13 we shall apply (58) to the simple situation of a differential current element in which I is a sinusoidal function of time. Other simple applications of (58) are considered in several problems at the end of this chapter.

We may summarize the use of the potentials by stating that a knowledge of the distribution of ρ_v and $\mathbf{J}$ throughout space theoretically enables us to determine V and $\mathbf{A}$ from (57) and (58). The electric and magnetic fields are then obtained by applying (50) and (51). If the charge and current distributions are unknown, or reasonable approximations cannot be made for them, these potentials usually offer no easier path toward the solution than does the direct application of Maxwell's equations.

✔ **D10.7.** A point charge of $4 \cos 10^8 \pi t \, \mu\text{C}$ is located at $P_+(0, 0, 1.5)$, while $-4 \cos 10^8 \pi t \, \mu\text{C}$ is at $P_-(0, 0, -1.5)$, both in free space. Find V at $P(r = 450, \theta, \phi = 0)$ at $t = 15$ ns for $\theta =$: (*a*) 0°; (*b*) 90°; (*c*) 45°.

Ans. 159.8 V; 0; 107.1 V

SUGGESTED REFERENCES

1. Bewley, L. V.: "Flux Linkages and Electromagnetic Induction," The Macmillan Company, New York, 1952. This little book discusses many of the paradoxical examples involving induced (?) voltages.
2. Faraday, M.: "Experimental Researches in Electricity," B. Quaritch, London, 1839, 1855. Very interesting reading of early scientific research. A more recent and available source is "Great Books of the Western World," vol. 45, Encyclopaedia Britannica, Inc., Chicago, 1952.

3. Halliday, D., R. Resnick and J. Walker: "Fundamentals of Physics," 5th ed. John Wiley & Sons, New York, 1997. This text is widely used in the first university-level course in physics.
4. Harman, W. W.: "Fundamentals of Electronic Motion," McGraw-Hill Book Company, New York, 1953. Relativistic effects are discussed in a clear and interesting manner.
5. Nussbaum, A.: "Electromagnetic Theory for Engineers and Scientists," Prentice-Hall, Inc., Englewood Cliffs, N.J., 1965. See the rocket-generator example beginning on p. 211.
6. Owen, G. E.: "Electromagnetic Theory," Allyn and Bacon, Inc., Boston, 1963. Faraday's law is discussed in terms of the frame of reference in chap. 8.
7. Panofsky, W. K. H. and M. Phillips: "Classical Electricity and Magnetism," 2d ed., Addison-Wesley Publishing Company, Inc., Reading, Mass., 1962. Relativity is treated at a moderately advanced level in chap. 15.

PROBLEMS

10.1 In Fig. 10.4, let $B = 0.2\cos 120\pi t$ T, and assume that the conductor joining the two ends of the resistor is perfect. It may be assumed that the magnetic field produced by $I(t)$ is negligible. Find: (*a*) $V_{ab}(t)$; (*b*) $I(t)$.

10.2 Given the time-varying magnetic field $\mathbf{B} = (0.5\mathbf{a}_x + 0.6\mathbf{a}_y - 0.3\mathbf{a}_z)\cos 5000t$ T and a square filamentary loop with its corners at (2,3,0), (2,−3,0), (−2,3,0), and (−2,−3,0), find the time-varying current flowing in the general $\mathbf{a}_\phi$ direction if the total loop resistance is 400 kΩ.

10.3 Given $\mathbf{H} = 300\mathbf{a}_z \cos(3 \times 10^8 t - y)$ A/m in free space, find the emf developed in the general $\mathbf{a}_\phi$ direction about the closed path having corners at: (*a*) (0,0,0), (1,0,0), (1,1,0), and (0,1,0); (*b*) (0,0,0) (2π,0,0), (2π,2π,0), (0, 2π,0).

10.4 Conductor surfaces are located at $\rho = 1$ cm and $\rho = 2$ cm in free space. The volume 1 cm $< \rho < 2$ cm contains the fields $H_\phi =$

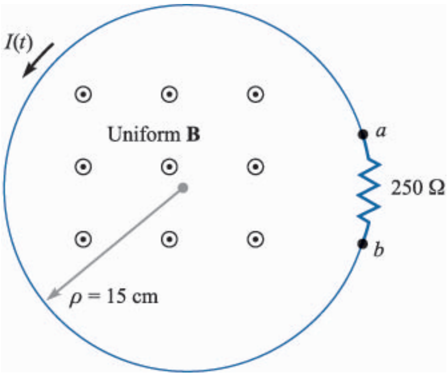


FIGURE 10.4
See Prob. 10.1.

$\frac{2}{\rho}\cos(6\times10^8\pi t - 2\pi z)$ A/m and $E_\rho = \frac{240\pi}{\rho}\cos(6\times10^8\pi t - 2\pi z)$ V/m. (*a*) Show that these two fields satisfy Eq. (6), Sec. 10.1. (*b*) Evaluate both integrals in Eq. (4) for the planar surface defined by $\phi = 0$, 1 cm $< \rho <$ 2 cm, $z = 0.1$, and its perimeter, and show that the same results are obtained.

10.5 The location of the sliding bar in Figure 10.5 is given by $x = 5t + 2t^3$, and the separation of the two rails is 20 cm. Let $\mathbf{B} = 0.8x^2\mathbf{a}_z$ T. Find the voltmeter reading at: (*a*) $t = 0.4$ s; (*b*) $x = 0.6$ m.

10.6 A perfectly conducting filament containing a small 500-Ω resistor is formed into a square, as illustrated by Fig. 10.6. Find $I(t)$ if $\mathbf{B} =$: (*a*) $0.3\cos(120\pi t - 30^\circ)\mathbf{a}_z$ T; (*b*) $0.4\cos[\pi(ct - y)]\mathbf{a}_z\ \mu$T, where $c = 3\times10^8$ m/s.

10.7 The rails in Fig. 10.7 each have a resistance of 2.2 Ω/m. The bar moves to the right at a constant speed of 9 m/s in a uniform magnetic field of 0.8 T. Find $I(t)$, $0 < t < 1$ s, if the bar is at $x = 2$ m at $t = 0$ and: (*a*) a 0.3-Ω resistor is present across the left end with the right end open-circuited; (*b*) a 0.3-Ω resistor is present across each end.

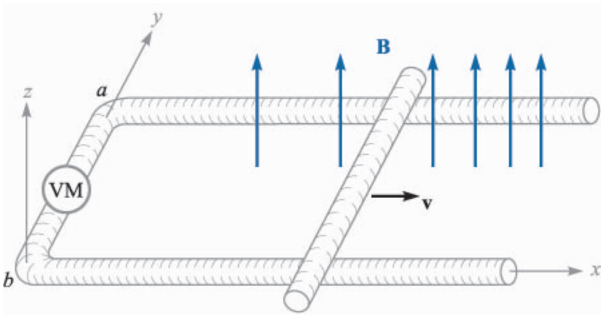


FIGURE 10.5
See Prob. 10.5.

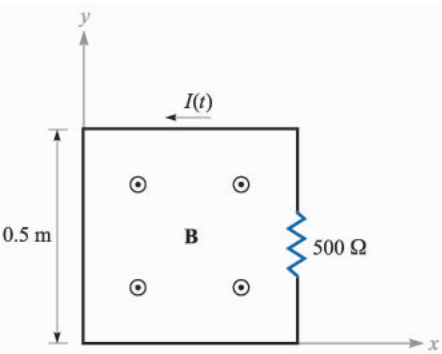


FIGURE 10.6
See Prob. 10.6.

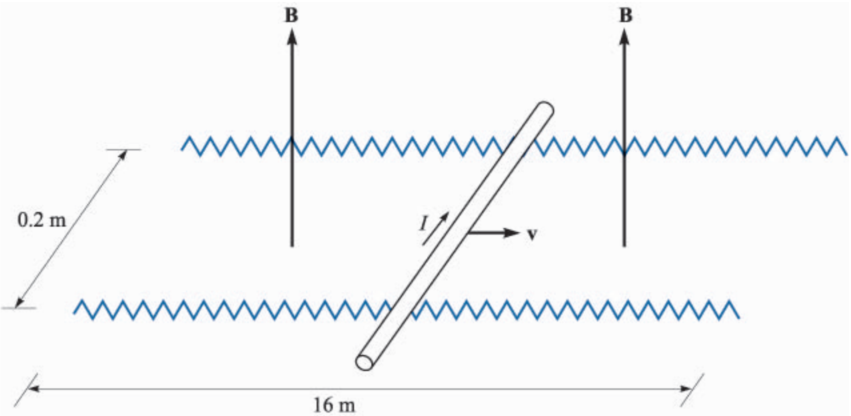


FIGURE 10.7
See Prob. 10.7.

10.8 Fig. 10.1 is modified to show that the rail separation is larger when y is larger. Specifically, let the separation $d = 0.2 + 0.02y$. Given a uniform velocity $v_y = 8$ m/s and a uniform magnetic flux density $B_z = 1.1$ T, find V_{12} as a function of time if the bar is located at $y = 0$ at $t = 0$.

10.9 A square filamentary loop of wire is 25 cm on a side and has a resistance of 125 Ω per meter length. The loop lies in the $z = 0$ plane with its corners at (0,0,0), (0.25,0,0), (0.25,0.25,0), and (0,0.25,0) at $t = 0$. The loop is moving with a velocity $v_y = 50$ m/s in the field $B_z = 8\cos(1.5 \times 10^8 t - 0.5x)\,\mu$T. Develop a function of time which expresses the ohmic power being delivered to the loop.

10.10 (*a*) Show that the ratio of the amplitudes of the conduction current density and the displacement current density is $\sigma/\omega\epsilon$ for the applied field $E = E_m \cos \omega t$. Assume $\mu = \mu_0$. (*b*) What is the amplitude ratio if the applied field is $E = E_m e^{-t/\tau}$, where τ is real?

10.11 Let the internal dimensions of a coaxial capacitor be $a = 1.2$ cm, $b = 4$ cm, and $l = 40$ cm. The homogeneous material inside the capacitor has the parameters $\epsilon = 10^{-11}$ F/m, $\mu = 10^{-5}$ H/m, and $\sigma = 10^{-5}$ S/m. If the electric field intensity is $\mathbf{E} = (10^6/\rho)\cos 10^5 \mathbf{a}_\rho$ V/m, find: (*a*) $\mathbf{J}$; (*b*) the total conduction current I_c through the capacitor; (*c*) the total displacement current I_d through the capacitor; (*d*) the ratio of the amplitude of I_d to that of I_c, the quality factor of the capacitor.

10.12 Given a coaxial transmission line with $\dfrac{b}{a} = e^{2.5}$, $\mu_R = \epsilon_R = 1$, and an electric field intensity $\mathbf{E} = \dfrac{200}{\rho}\cos(10^9 t - 3.336z)\mathbf{a}_\rho$ V/m, find: (*a*) V_{ab}, the voltage between the conductors, if it is known that the electrostatic relationship $\mathbf{E} = -\nabla V$ is valid; (*b*) the displacement current density.

10.13 Consider the region defined by $|x|$, $|y|$, and $|z| < 1$. Let $\epsilon_R = 5$, $\mu_R = 4$, and $\sigma = 0$. If $J_d = 20\cos(1.5 \times 10^8 t - bx)\mathbf{a}_y$ μA/m^2: (*a*) find **D** and **E**; (*b*) use the point form of Faraday's law and an integration with respect to time to find **B** and **H**; (*c*) use $\nabla \times \mathbf{H} = \mathbf{J}_d + \mathbf{J}$ to find $\mathbf{J}_d$. (*d*) What is the numerical value of b?

10.14 A voltage source $V_0 \sin \omega t$ is connected between two concentric conducting spheres, $r = a$ and $r = b$, $b > a$, where the region between them is a material for which $\epsilon = \epsilon_R \epsilon_0$, $\mu = \mu_0$, and $\sigma = 0$. Find the total displacement current through the dielectric and compare it with the source current as determined from the capacitance (Sec. 5.10) and circuit-analysis methods.

10.15 Let $\mu = 3 \times 10^{-5}$ H/m, $\epsilon = 1.2 \times 10^{-10}$ F/m, and $\sigma = 0$ everywhere. If $\mathbf{H} = 2\cos(10^{10} t - \beta x)\mathbf{a}_z$ A/m, use Maxwell's equations to obtain expressions for **B**, **D**, **E**, and β.

10.16 (*a*) A certain material has $\sigma = 0$ and $\epsilon_R = 1$. If $\mathbf{H} = 4\sin(10^6 t - 0.01z)\mathbf{a}_y$ A/m, make use of Maxwell's equations to find μ_R. (*b*) Find $\mathbf{E}(z, t)$.

10.17 The electric field intensity in the region $0 < x < 5$, $0 < y < \pi/12$, $0 < z < 0.06$ m in free space is given by $\mathbf{E} = C \sin 12y \sin az \cos 2 \times 10^{10} t\mathbf{a}_x$ V/m. Beginning with the $\nabla \times \mathbf{E}$ relationship, use Maxwell's equations to find a numerical value for a, if it is known that a is greater than zero.

10.18 The parallel-plate transmission line shown in Fig. 10.8 has dimensions $b = 4$ cm and $d = 8$ mm, while the medium between the plates is characterized by $\mu_R = 1$, $\epsilon_R = 20$, and $\sigma = 0$. Neglect fields outside the dielectric. Given the field $\mathbf{H} = 5\cos(10^9 t - \beta z)\mathbf{a}_y$ A/m, use Maxwell's equations to help find: (*a*) β, if $\beta > 0$; (*b*) the displacement current density at $z = 0$; (*c*) the total displacement current crossing the surface $x = 0.5d$, $0 < y < b$, $0 < z < 0.1$ m in the $\mathbf{a}_x$ direction.

10.19 In the first section of this chapter, Faraday's law was used to show that the field $\mathbf{E} = -\frac{1}{2}kB_0 e^{kt}\mathbf{a}_\phi$ results from the changing magnetic field $\mathbf{B} = B_0 e^{kt}\mathbf{a}_z$. (*a*) Show that these fields do not satisfy Maxwell's other

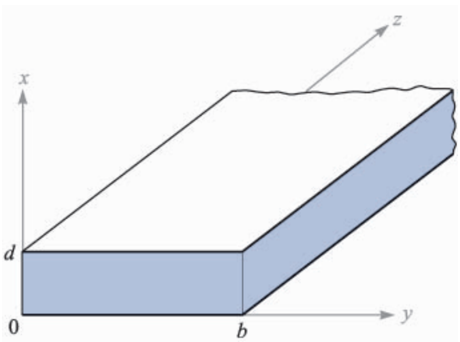


FIGURE 10.8
See Prob. 10.18.

curl equation. (*b*) If we let $B_0 = 1$ T and $k = 10^6\ s^{-1}$, we are establishing a fairly large magnetic flux density in 1 μs. Use the $\nabla \times \mathbf{H}$ equation to show that the rate at which B_z should (but does not) change with ρ is only about 5×10^{-6} T per meter in free space at $t = 0$.

10.20 Point $C(-0.1, -0.2, 0.3)$ lies on the surface of a perfect conductor. The electric field intensity at C is $(500\mathbf{a}_x - 300\mathbf{a}_y + 600\mathbf{a}_z)\cos 10^7 t$ V/m, and the medium surrounding the conductor is characterized by $\mu_R = 5$, $\epsilon_R = 10$, and $\sigma = 0$. (*a*) Find a unit vector normal to the conductor surface at C, if the origin lies within the conductor. (*b*) Find the surface charge density at C.

10.21 The surfaces $\rho = 3$ and 10 mm, and $z = 0$ and 25 cm are perfect conductors. The region enclosed by these surfaces has $\mu = 2.5 \times 10^{-6}$ H/m, $\epsilon = 4 \times 10^{-11}$ F/m, and $\sigma = 0$. Let $\mathbf{H} = (2/\rho)\cos 10\pi z \cos \omega t \mathbf{a}_\phi$ A/m. Make use of Maxwell's equations to find: (*a*) ω; (*b*) $\mathbf{E}$.

10.22 In free space, where $\epsilon = \epsilon_0$, $\mu = \mu_0$, $\sigma = 0$, $\mathbf{J} = 0$, and $\rho_v = 0$, assume a cartesian coordinate system in which $\mathbf{E}$ and $\mathbf{H}$ are both functions only of z and t. (*a*) If $\mathbf{E} = E_y\mathbf{a}_y$ and $\mathbf{H} = H_x\mathbf{a}_x$, begin with Maxwell's equations and determine the second-order partial differential equation that E_y must satisfy. (*b*) Show that $E_y = 5(300t + bz)^2$ is a solution of that equation for a particular value of b, and find that value.

10.23 In region 1, $z < 0$, $\epsilon_1 = 2 \times 10^{-11}$ F/m, $\mu_1 = 2 \times 10^{-6}$ H/m, and $\sigma_1 = 4 \times 10^{-3}$ S/m; in region 2, $z > 0$, $\epsilon_2 = \epsilon_1/2$, $\mu_2 = 2\mu_1$, and $\sigma_2 = \sigma_1/4$. It is known that $\mathbf{E}_1 = (30\mathbf{a}_x + 20\mathbf{a}_y + 10\mathbf{a}_z)\cos 10^9 t$ V/m at $P(0,0,0^-)$. (*a*) Find $\mathbf{E}_{N1}$, $\mathbf{E}_{t1}$, $\mathbf{D}_{N1}$, and $\mathbf{D}_{t1}$ at P_1. (*b*) Find $\mathbf{J}_{N1}$ and $\mathbf{J}_{t1}$ at P_1. (*c*) Find $\mathbf{E}_{t2}$, $\mathbf{D}_{t2}$, and $\mathbf{J}_{t2}$ at $P_2(0, 0, 0^+)$. (*d*) (Harder) Use the continuity equation to help show that $J_{N1} - J_{N2} = \partial D_{N2}/\partial t$, and then determine $\mathbf{D}_{N2}$, $\mathbf{J}_{N2}$, and $\mathbf{E}_{N2}$.

10.24 Given the fields $V = 80z \cos x \cos 3 \times 10^8 t$ kV and $\mathbf{A} = 26.7z \sin x \sin 3 \times 10^8 t\mathbf{a}_x$ mWb/m in free space, find $\mathbf{E}$ and $\mathbf{H}$.

10.25 In a region where $\mu_R = \epsilon_R = 1$ and $\sigma = 0$, the retarded potentials are given by $V = x(z - ct)$ V and $\mathbf{A} = x\left(\frac{z}{c} - t\right)\mathbf{a}_z$ Wb/m, where $c = 1\sqrt{\mu_0\epsilon_0}$. (*a*) Show that $\nabla \cdot \mathbf{A} = -\mu\epsilon \frac{\partial V}{\partial t}$. (*b*) Find $\mathbf{B}$, $\mathbf{H}$, $\mathbf{E}$, and $\mathbf{D}$. (*c*) Show that these results satisfy Maxwell's equations if $\mathbf{J}$ and ρ_v are zero.

10.26 Let the current $I = 80t$ A be present in the $\mathbf{a}_z$ direction on the z axis in free space within the interval $-0.1 < z < 0.1$ m. (*a*) Find A_z at $P(0, 2, 0)$ and: (*b*) sketch A_z versus t over the time interval $-0.1 < t < 0.1$ μs.

CHAPTER

11

THE UNIFORM PLANE WAVE

In this chapter we shall apply Maxwell's equations to introduce the fundamental theory of wave motion. The uniform plane represents one of the simplest applications of Maxwell's equations, and yet it is of profound importance, since it is a basic entity by which energy is propagated. We shall explore the physical processes that determine the speed of propagation and the extent to which attenuation may occur. We shall derive and make use of the Poynting theorem to find the power carried by a wave. Finally, we shall learn how to describe wave polarization. This chapter is the foundation for our explorations in later chapters which will include wave reflection, basic transmission line and waveguiding concepts, and wave generation through antennas.

11.1 WAVE PROPAGATION IN FREE SPACE

As we indicated in our discussion of boundary conditions in the previous chapter, the solution of Maxwell's equations without the application of any boundary conditions at all represents a very special type of problem. Although we restrict our attention to a solution in rectangular coordinates, it may seem even then that we are solving several different problems as we consider various special cases in this chapter. Solutions are obtained first for free-space conditions, then for perfect dielectrics, next for lossy dielectrics, and finally for the good conductor. We do this to take advantage of the approximations that are applicable to each

special case and to emphasize the special characteristics of wave propagation in these media, but it is not necessary to use a separate treatment; it is possible (and not very difficult) to solve the general problem once and for all.

To consider wave motion in free space first, Maxwell's equations may be written in terms of **E** and **H** only as

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \tag{1}$$

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \tag{2}$$

$$\nabla \cdot \mathbf{E} = 0 \tag{3}$$

$$\nabla \cdot \mathbf{H} = 0 \tag{4}$$

Now let us see whether wave motion can be inferred from these four equations without actually solving them. The first equation states that if **E** is changing with time at some point, then **H** has curl at that point and thus can be considered as forming a small closed loop linking the changing **E** field. Also, if **E** is changing with time, then **H** will in general also change with time, although not necessarily in the same way. Next, we see from the second equation that this changing **H** produces an electric field which forms small closed loops about the **H** lines. We now have once more a changing electric field, our original hypothesis, but this field is present a small distance away from the point of the original disturbance. We might guess (correctly) that the velocity with which the effect moves away from the original point is the velocity of light, but this must be checked by a more quantitative examination of Maxwell's equations.

Let us first write Maxwell's four equations above for the special case of sinusoidal (more strictly, cosinusoidal) variation with time. This is accomplished by complex notation and phasors. The procedure is identical to the one we used in studying the sinusoidal steady state in electric circuit theory.

Given the vector field

$$\mathbf{E} = E_x \mathbf{a_x}$$

we assume that the component E_x is given as

$$E_x = E(x, y, z)\cos(\omega t + \psi) \tag{5}$$

where $E(x, y, z)$ is a real function of x, y, z and perhaps ω, but not of time, and ψ is a phase angle which may also be a function of x, y, z and ω. Making use of Euler's identity,

$$e^{j\omega t} = \cos \omega t + j \sin \omega t$$

we let

$$E_x + \text{Re}[E(x, y, z)e^{j(\omega t+\psi)}] = \text{Re}[E(x, y, z)e^{j\psi}e^{j\omega t}] \tag{6}$$

where Re signifies that the real part of the following quantity is to be taken. If we then simplify the nomenclature by dropping Re and suppressing $e^{j\omega t}$, the field quantity E_x becomes a phasor, or a complex quantity, which we identify by use of an s subscript, E_{xs}. Thus

$$E_{xs} = E(x, y, z)e^{j\psi} \tag{7}$$

and

$$\mathbf{E_s} = E_{xs}\mathbf{a_x}$$

The s can be thought of as indicating a frequency domain quantity expressed as a function of the complex frequency s, even though we shall consider only those cases in which s is a pure imaginary, $s = j\omega$.

Example 11.1

Let us express $E_y = 100\cos(10^8 t - 0.5z + 30^\circ)$ V/m as a phasor.

Solution. We first go to exponential notation,

$$E_y = \text{Re}[100e^{j(10^8 t - 0.5z + 30^\circ)}]$$

and then drop Re and suppress $e^{j10^8 t}$, obtaining the phasor

$$E_{ys} = 100e^{-j0.5z + j30^\circ}$$

Note that E_y is real, but E_{ys} is in general complex. Note also that a mixed nomenclature is commonly used for the angle. That is, $0.5z$ is in radians, while 30° is in degrees.

Given a scalar component or a vector expressed as a phasor, we may easily recover the time-domain expression.

Example 11.2

Given the field intensity vector, $\mathbf{E_s} = 100\angle 30^\circ \mathbf{a_x} + 20\angle -50^\circ \mathbf{a_y} + 40\angle 210^\circ \mathbf{a_z}$ V/m, identified as a phasor by its subscript s, we desire the vector as a real function of time.

Solution. Our starting point is the phasor,

$$\mathbf{E_s} = 100\angle 30^\circ \mathbf{a_x} + 20\angle -50^\circ \mathbf{a_y} + 40\angle 210^\circ \mathbf{a_z} \text{ V/m}$$

Let us assume that the frequency is specified as 1 MHz. We first select exponential notation for mathematical clarity,

$$\mathbf{E_s} = 100e^{j30^\circ}\mathbf{a_x} + 20e^{-j50^\circ}\mathbf{a_y} + 40e^{j210^\circ}\mathbf{a_z} \text{ V/m}$$

reinsert the $e^{j\omega t}$ factor,

$$\mathbf{E_s(t)} = (100e^{j30^\circ}\mathbf{a_x} + 20e^{-j50^\circ}\mathbf{a_y} + 40e^{j210^\circ}\mathbf{a_z})e^{j2\pi 10^6 t}$$
$$= 100e^{j(2\pi 10^6 t + 30^\circ)}\mathbf{a_x} + 20e^{j(2\pi 10^6 t - 50^\circ)}\mathbf{a_y} + 40e^{j(2\pi 10^6 t + 210^\circ)}\mathbf{a_z}$$

and take the real part, obtaining the real vector,

$$\mathbf{E(t)} = 100\cos(2\pi 10^6 t + 30^\circ)\mathbf{a_x} + 20\cos(2\pi 10^6 t - 50^\circ)\mathbf{a_y} + 40\cos(2\pi 10^6 t + 210^\circ)\mathbf{a_z}$$

None of the amplitudes or phase angles in this example are expressed as a function of x, y, or z, but, if any are, the same procedure is effective. Thus, if $\mathbf{H_s} = 20e^{-(0.1+j20)z}\mathbf{a_x}$ A/m, then

$$\mathbf{H(t)} = \text{Re}[20e^{-0.1z}e^{-j20z}e^{j\omega t}]\mathbf{a_x} = 20e^{-0.1z}\cos(\omega t - 20z)\mathbf{a_x} \text{ A/m}$$

Now, since

$$\frac{\partial E_x}{\partial t} = \frac{\partial}{\partial t}[E(x, y, z)\cos(\omega t + \psi)] = -\omega E(x, y, z)\sin(\omega t + \psi)$$
$$= Re[j\omega E_{xs} e^{j\omega t}]$$

it is evident that taking the partial derivative of any field quantity with respect to time is equivalent to multiplying the corresponding phasor by $j\omega$. As an example, if

$$\frac{\partial E_x}{\partial t} = -\frac{1}{\epsilon_0}\frac{\partial H_y}{\partial z}$$

the corresponding phasor expression is

$$j\omega E_{xs} = -\frac{1}{\epsilon_0}\frac{\partial H_{ys}}{\partial z}$$

where E_{xs} and H_{ys} are complex quantities. We next apply this notation to Maxwell's equations. Thus, given the equation,

$$\nabla \times \mathbf{H} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

the corresponding relationship in terms of phasor-vectors is

$$\boxed{\nabla \times \mathbf{H_s} = j\omega\epsilon_0 \mathbf{E_s}} \tag{8}$$

Equation (8) and the three equations

$$\nabla \times \mathbf{E_s} = -j\omega\mu_0 \mathbf{H_s} \tag{9}$$
$$\nabla \cdot \mathbf{E_s} = 0 \tag{10}$$
$$\nabla \cdot \mathbf{H_s} = 0 \tag{11}$$

are Maxwell's four equations in phasor notation for sinusoidal time variation in free space. It should be noted that (10) and (11) are no longer independent relationships, for they can be obtained by taking the divergence of (8) and (9), respectively.

Our next step is to obtain the sinusoidal steady-state form of the wave equation, a step we could omit because the simple problem we are going to solve yields easily to simultaneous solution of the four equations above. The wave equation is an important equation, however, and it is a convenient starting point for many other investigations.

The method by which the wave equation is obtained could be accomplished in one line (using four equals signs on a wider sheet of paper):

$$\nabla \times \nabla \times \mathbf{E_s} = \nabla(\nabla \cdot \mathbf{E_s}) - \nabla^2 \mathbf{E_s} = -j\omega\mu_0 \nabla \times \mathbf{H_s}$$
$$= \omega^2 \mu_0 \epsilon_0 \mathbf{E_s} = -\nabla^2 \mathbf{E_s}$$

since $\nabla \cdot \mathbf{E_s} = 0$. Thus

$$\boxed{\nabla^2 \mathbf{E_s} = -k_0^2 \mathbf{E_s}} \tag{12}$$

where k_0, the free space *wavenumber*, is defined as

$$\boxed{k_0 = \omega\sqrt{\mu_0 \epsilon_0}} \tag{13}$$

Eq. (12) is known as the vector Helmholtz equation.[1] It is fairly formidable when expanded, even in rectangular coordinates, because three scalar phasor equations result, and each has four terms. The x component of (12) becomes, still using the del-operator notation,

$$\nabla^2 \mathbf{E_{xs}} = -k_0^2 \mathbf{E_{xs}} \tag{14}$$

and the expansion of the operator leads to the second-order partial differential equation

$$\frac{\partial^2 E_{xs}}{\partial x^2} + \frac{\partial^2 E_{xs}}{\partial y^2} + \frac{\partial^2 E_{xs}}{\partial z^2} = -k_0^2 E_{xs} \tag{15}$$

Let us attempt a solution of (15) by assuming that a simple solution is possible in which E_{xs} does not vary with x or y, so that the two corresponding derivatives are zero, leading to the ordinary differential equation

$$\frac{d^2 E_{xs}}{dz^2} = -k_0^2 E_{xs} \tag{16}$$

By inspection, we may write down one solution of (16):

$$E_{xs} = E_{x0} e^{-jk_0 z} \tag{17}$$

[1] Hermann Ludwig Ferdinand von Helmholtz (1821–1894) was a professor at Berlin working in the fields of physiology, electrodynamics, and optics. Hertz was one of his students.

Next, we reinsert the $e^{j\omega t}$ factor and take the real part,

$$\boxed{E_x(z, t) = E_{x0}\cos(\omega t - k_0 z)} \tag{18}$$

where the amplitude factor, E_{x0}, is the value of E_x at $z = 0$, $t = 0$. Problem 1 at the end of the chapter indicates that

$$\boxed{E'_x(z, t) = E'_{x0}\cos(\omega t + k_0 z)} \tag{19}$$

may also be obtained from an alternate solution of the vector Helmholtz equation.

We refer to the solutions expressed in (18) and (19) as the *real instantaneous* forms of the electric field. They are the mathematical representations of what one would experimentally measure. The terms ωt and $k_0 z$, appearing in (18) and (19), have units of angle, and are usually expressed in radians. We know that ω is the radian time frequency, measuring phase shift *per unit time*, and which has units of *rad/sec*. In a similar way, we see that k_0 will be interpreted as a *spatial* frequency, which in the present case measures the phase shift *per unit distance* along the z direction. Its units are rad/m. In addition to its original name (free space wavenumber), k_0 is also the *phase constant* for a uniform plane wave in free space.

We see that the fields of (18) and (19) are x components, which we might describe as directed upward at the surface of a plane earth. The radical $\sqrt{\mu_0\epsilon_0}$, contained in k_0, has the approximate value $1/(3 \times 10^8)$ s/m, which is the reciprocal of c, the velocity of light in free space,

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} = 2.998 \times 10^8 \doteq 3 \times 10^8 \text{ m/s}$$

We can thus write $k_0 = \omega/c$, and Eq. (18), for example, can be rewritten as

$$E_x(z, t) = E_{x0}\cos[\omega(t - z/c)] \tag{20}$$

The propagation wave nature of the fields as expressed in (18), (19), and (20) can now be seen. First, suppose we were to fix the time at $t = 0$. Eq. (20) then becomes

$$E_x(z, 0) = E_{x0}\cos\left(\frac{\omega z}{c}\right) = E_{x0}\cos(k_0 z) \tag{21}$$

which we identify as a simple periodic function that repeats every incremental distance λ, known as the *wavelength*. The requirement is that $k_0\lambda = 2\pi$, and so

$$\boxed{\lambda = \frac{2\pi}{k_0} = \frac{c}{f} = \frac{3 \times 10^8}{f} \qquad \text{(free space)}} \tag{22}$$

Now suppose we consider some point (such as a wave crest) on the cosine function of Eq. (21). For a crest to occur, the argument of the cosine must be an integer multiple of 2π. Considering the mth crest of the wave, the condition becomes

$$k_0 z = 2m\pi$$

So let us now consider the point on the cosine that we have chosen, and see what happens as time is allowed to increase. Eq. (18) now applies, where our requirement is that the entire cosine argument be the same multiple of 2π for all time, in order to keep track of the chosen point. From (18) and (20) our condition now becomes

$$\omega t - k_0 z = \omega(t - z/c) = 2m\pi \tag{23}$$

We see that as time increases (as it must), the position z must also increase in order to satisfy (23). *Thus the wave crest (and the entire wave) moves in the positive z direction*. The speed of travel, or wave *phase velocity*, is given by c (in free space), as can be deduced from (23). Using similar reasoning, Eq. (19), having cosine argument $(\omega t + k_0 z)$, describes a wave that moves in the *negative z* direction, since as time increases, z must now *decrease* to keep the argument constant. Waves expressed in the forms exemplified by Eqs. (18) and (19) are called *traveling waves*. For simplicity, we will restrict our attention in this chapter to only the positive z traveling wave.

Let us now return to Maxwell's equations, (8) to (11), and determine the form of the **H** field. Given $\mathbf{E_s}$, $\mathbf{H_s}$ is most easily obtained from (9),

$$\nabla \times \mathbf{E_s} = -j\omega\mu_0 \mathbf{H_s} \tag{9}$$

which is greatly simplified for a single $\mathbf{E_{xs}}$ component varying only with z,

$$\frac{dE_{xs}}{dz} = -j\omega\mu_0 H_{ys}$$

Using (17) for E_{xs}, we have

$$H_{ys} = -\frac{1}{j\omega\mu_0}(-jk_0)E_{x0}e^{-jk_0 z} = E_{x0}\sqrt{\frac{\epsilon_0}{\mu_0}}e^{-jk_0 z}$$

In real instantaneous form, this becomes

$$\boxed{H_y(z, t) = E_{x0}\sqrt{\frac{\epsilon_0}{\mu_0}}\cos(\omega t - k_0 z)} \tag{24}$$

where E_{x0} is assumed real.

We therefore find the x-directed **E** field that propagates in the positive z direction is accompanied by a y-directed **H** field. Moreover, the ratio of the electric and magnetic field intensities, given by the ratio of (18) to (24),

$$\frac{E_x}{H_y} = \sqrt{\frac{\mu_0}{\epsilon_0}} \tag{25}$$

is constant. Using the language of circuit theory, we would say that E_x and H_y are "in phase," but this in-phase relationship refers to space as well as to time. We are accustomed to taking this for granted in a circuit problem in which a

current $I_m \cos \omega t$ is assumed to have its maximum amplitude I_m throughout an entire series circuit at $t = 0$. Both (18) and (24) clearly show, however, that the maximum value of either E_x or H_y occurs when $\omega(t - z/c)$ is an integral multiple of 2π rad; neither field is a maximum everywhere at the same instant. It is remarkable, then, that the ratio of these two components, both changing in space and time, should be everywhere a constant.

The square root of the ratio of the permeability to the permittivity is called the *intrinsic impedance* η (eta),

$$\boxed{\eta = \sqrt{\frac{\mu}{\epsilon}}} \tag{26}$$

where η has the dimension of ohms. The intrinsic impedance of free space is

$$\boxed{\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = 377 \doteq 120\pi\ \Omega}$$

This wave is called a *uniform plane wave* because its value is uniform throughout any plane, $z =$ constant. It represents an energy flow in the positive z direction. Both the electric and magnetic fields are perpendicular to the direction of propagation, or both lie in a plane that is transverse to the direction of propagation; the uniform plane wave is a *transverse electromagnetic wave*, or a TEM wave.

Some feeling for the way in which the fields vary in space may be obtained from Figs 11.1*a* and 11.1*b*. The electric field intensity in Fig. 11.1*a* is shown at $t = 0$, and the instantaneous value of the field is depicted along three lines, the z axis and arbitrary lines parallel to the z axis in the $x = 0$ and $y = 0$ planes. Since the field is uniform in planes perpendicular to the z axis, the variation along all three of the lines is the same. One complete cycle of the variation occurs in a wavelength, λ. The values of H_y at the same time and positions are shown in Fig. 11.1*b*.

A uniform plane wave cannot exist physically, for it extends to infinity in two dimensions at least and represents an infinite amount of energy. The distant field of a transmitting antenna, however, is essentially a uniform plane wave in some limited region; for example, a radar signal impinging on a distant target is closely a uniform plane wave.

Although we have considered only a wave varying sinusoidally in time and space, a suitable combination of solutions to the wave equation may be made to achieve a wave of any desired form. The summation of an infinite number of harmonics through the use of a Fourier series can produce a periodic wave of square or triangular shape in both space and time. Nonperiodic waves may be obtained from our basic solution by Fourier integral methods. These topics are among those considered in the more advanced books on electromagnetic theory.

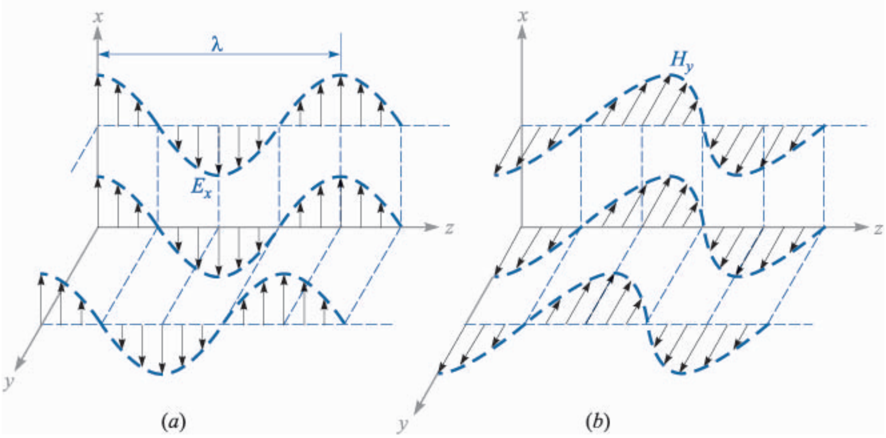


FIGURE 11.1
(*a*) Arrows represent the instantaneous values of $E_{x0}\cos[\omega(t - z/c)]$ at $t = 0$ along the z axis, along an arbitrary line in the $x = 0$ plane parallel to the z axis, and along an arbitrary line in the $y = 0$ plane parallel to the z axis. (*b*) Corresponding values of H_y are indicated. Note that E_x and H_y are in phase at any point at any time.

✔ **D11.1.** The electric field amplitude of a uniform plane wave propagating in the $\mathbf{a_z}$ direction is 250 V/m. If $\mathbf{E} = E_x\mathbf{a_x}$ and $\omega = 1.00$ Mrad/s, find: (a) the frequency; (b) the wavelength; (c) the period; (d) the amplitude of **H**.

Ans. 159 kHz; 1.88 km; 6.28 µs; 0.663 A/m

✔ **D11.2** Let $\mathbf{H_s} = (2\angle -40^\circ \mathbf{a_x} - 3\angle 20^\circ \mathbf{a_y})e^{-j0.07z}$ A/m for a uniform plane wave traveling in free space. Find: (a) ω; (b) H_x at $P(1, 2, 3)$ at $t = 31$ ns; (c) $|\mathbf{H}|$ at $t = 0$ at the origin.

Ans. 21.0 Mrad/s; 1.93 A/m; 3.22 A/m

11.2 WAVE PROPAGATION IN DIELECTRICS

Let us now extend our analytical treatment of the uniform plane wave to propagation in a dielectric of permittivity ϵ and permeability μ. The medium is isotropic and homogeneous, and the wave equation is now

$$\boxed{\nabla^2 \mathbf{E_s} = -k^2 \mathbf{E_s}} \qquad (27)$$

where the wavenumber is now a function of the material properties:

$$\boxed{k = \omega\sqrt{\mu\epsilon} = k_0\sqrt{\mu_R\epsilon_R}} \qquad (28)$$

For E_{xs} we have

$$\frac{d^2 E_{xs}}{dz^2} = -k^2 E_{xs} \tag{29}$$

An important feature of wave propagation in a dielectric is that k can be complex-valued, and as such is referred to as the complex *propagation constant*. A general solution of (29) in fact allows the possibility of a complex k, and it is customary to write it in terms of its real and imaginary parts in the following way:

$$\boxed{jk = \alpha + j\beta} \tag{30}$$

A solution of (29) will be:

$$\boxed{E_{xs} = E_{x0}e^{-jkz} = E_{x0}e^{-\alpha z}e^{-j\beta z}} \tag{31}$$

Multiplying (31) by $e^{j\omega t}$ and taking the real part yields a form of the field that can be more easily visualized:

$$\boxed{E_x = E_{x0}e^{-\alpha z}\cos(\omega t - \beta z)} \tag{32}$$

We recognize the above as a uniform plane wave that propagates in the forward z direction with phase constant β, but which (for positive α) loses amplitude with increasing z according to the factor $e^{-\alpha z}$. *Thus the general effect of a complex-valued k is to yield a traveling wave that changes its amplitude with distance*. If α is positive, it is called the *attenuation coefficient*. If α is negative, the wave grows in amplitude with distance, and α is called the *gain coefficient*. The latter effect would occur, for example, in laser amplifiers. In the present and future discussions in this book, we will consider only passive media, in which one or more loss mechanisms are present, thus producing a positive α.

The attenuation coefficient is measured in nepers per meter (Np/m) in order that the exponent of e be measured in the dimensionless units of nepers.[2] Thus, if $\alpha = 0.01$ Np/m, the crest amplitude of the wave at $z = 50$ m will be $e^{-0.5}/e^{-0} = 0.607$ of its value at $z = 0$. In traveling a distance $1/\alpha$ in the $+z$ direction, the amplitude of the wave is reduced by the familiar factor of e^{-1}, or 0.368.

[2] The term *neper* was selected (by some poor speller) to honor John Napier, a Scottish mathematician who first proposed the use of logarithms.

The ways in which physical processes in a material can affect the wave electric field are described through a *complex permittivity* of the form

$$\epsilon = \epsilon' - j\epsilon'' \tag{33}$$

Two important mechanisms that give rise to a complex permittivity (and thus result in wave losses) are bound electron or ion oscillations and dipole relaxation, both of which are discussed in Appendix D. An additional mechanism is the conduction of free electrons or holes, which we will explore at length in this chapter.

Losses arising from the response of the medium to the magnetic field can occur as well, and are modeled through a *complex permeability*, $\mu = \mu' - j\mu''$. Examples of such media include *ferrimagnetic* materials, or *ferrites*. The magnetic response is usually very weak compared to the dielectric response in most materials of interest for wave propagation; in such materials $\mu \approx \mu_0$. Consequently, our discussion of loss mechanisms will be confined to those described through the complex permittivity.

We can substitute (33) into (28), which results in

$$k = \omega\sqrt{\mu(\epsilon' - j\epsilon'')} = \omega\sqrt{\mu\epsilon'}\sqrt{1 - j\frac{\epsilon''}{\epsilon'}} \tag{34}$$

Note the presence of the second radical factor in (34), which becomes unity (and real) as ϵ'' vanishes. With non-zero ϵ'', k is complex, and so losses occur which are quantified through the attenuation coefficient, α, in (30). The phase constant, β (and consequently the wavelength and phase velocity), will also be affected by ϵ''. α and β are found by taking the real and imaginary parts of jk from (34). We obtain:

$$\alpha = \text{Re}\{jk\} = \omega\sqrt{\frac{\mu\epsilon'}{2}}\left(\sqrt{1 + \left(\frac{\epsilon''}{\epsilon'}\right)^2} - 1\right)^{1/2} \tag{35}$$

$$\beta = \text{Im}\{jk\} = \omega\sqrt{\frac{\mu\epsilon'}{2}}\left(\sqrt{1 + \left(\frac{\epsilon''}{\epsilon'}\right)^2} + 1\right)^{1/2} \tag{36}$$

We see that a non-zero α (and hence loss) results if the imaginary part of the permittivity, ϵ'', is present. We also observe in (35) and (36) the presence of the ratio ϵ''/ϵ', which is called the *loss tangent*. The meaning of the term will be demonstrated when we investigate the specific case of conductive media. The practical importance of the ratio lies in its magnitude compared to unity, which enables simplifications to be made in (35) and (36).

Whether or not losses occur, we see from (32) that the wave phase velocity is given by

$$v_p = \frac{\omega}{\beta} \tag{37}$$

The wavelength is the distance required to effect a phase change of 2π radians

$$\beta\lambda = 2\pi$$

which leads to the fundamental definition of wavelength,

$$\lambda = \frac{2\pi}{\beta} \tag{38}$$

Since we have a uniform plane wave, the magnetic field is found through

$$H_{ys} = \frac{E_{x0}}{\eta} e^{-\alpha z} e^{-j\beta z}$$

where the intrinsic impedance is now a complex quantity,

$$\eta = \sqrt{\frac{\mu}{\epsilon' - j\epsilon''}} = \sqrt{\frac{\mu}{\epsilon'}} \frac{1}{\sqrt{1 - j(\epsilon''/\epsilon')}} \tag{39}$$

The electric and magnetic fields are no longer in phase.

A special case is that of a lossless medium, or *perfect dielectric*, in which $\epsilon'' = 0$, and so $\epsilon = \epsilon'$. From (35), this leads to $\alpha = 0$, and from (36),

$$\beta = \omega\sqrt{\mu\epsilon'} = \omega\sqrt{\mu\epsilon} \qquad \text{(lossless medium)} \tag{40}$$

With $\alpha = 0$, the real field assumes the form:

$$E_x = E_{x0} \cos(\omega t - \beta z) \tag{41}$$

We may interpret this as a wave traveling in the $+z$ direction at a phase velocity v_p, where

$$v_p = \frac{\omega}{\beta} = \frac{1}{\sqrt{\mu\epsilon}} = \frac{c}{\sqrt{\mu_R \epsilon_R}}$$

The wavelength is

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{\omega\sqrt{\mu\epsilon}} = \frac{1}{f\sqrt{\mu\epsilon}} = \frac{c}{f\sqrt{\mu_R\epsilon_R}} = \frac{\lambda_0}{\sqrt{\mu_R\epsilon_R}} \qquad \text{(lossless medium)} \tag{42}$$

where λ_0 is the free space wavelength. Note that $\mu_R\epsilon_R > 1$, and therefore the wavelength is shorter and the velocity is lower in all real media than they are in free space.

Associated with E_x is the magnetic field intensity

$$H_y = \frac{E_{x0}}{\eta}\cos(\omega t - \beta z)$$

where the intrinsic impedance is

$$\eta = \sqrt{\frac{\mu}{\epsilon}} \tag{43}$$

The two fields are once again perpendicular to each other, perpendicular to the direction of propagation, and in phase with each other everywhere. Note that when **E** is crossed into **H**, the resultant vector is in the direction of propagation. We shall see the reason for this when we discuss the Poynting vector.

Example 11.3

Let us apply these results to a 1 MHz plane wave propagating in fresh water. At this frequency, losses in water are known to be small, so for simplicity, we will neglect ϵ''. In water, $\mu_R = 1$ and at 1 MHz, $\epsilon'_R = \epsilon_R = 81$.

Solution. We begin by calculating the phase constant. Using (36) with $\epsilon'' = 0$, we have

$$\beta = \omega\sqrt{\mu\epsilon'} = \omega\sqrt{\mu_0\epsilon_0}\sqrt{\epsilon'_R} = \frac{\omega\sqrt{\epsilon'_R}}{c} = \frac{2\pi \times 10^6\sqrt{81}}{3.0 \times 10^8} = 0.19 \text{ rad/m}$$

Using this result, we can determine the wavelength and phase velocity:

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{.19} = 33 \text{ m}$$

$$v_p = \frac{\omega}{\beta} = \frac{2\pi \times 10^6}{.19} = 3.3 \times 10^7 \text{ m/s}$$

The wavelength in air would have been 300 m. Continuing our calculations, we find the intrinsic impedance, using (39) with $\epsilon'' = 0$:

$$\eta = \sqrt{\frac{\mu}{\epsilon'}} = \frac{\eta_0}{\sqrt{\epsilon'_R}} = \frac{377}{9} = 42\ \Omega$$

If we let the electric field intensity have a maximum amplitude of 0.1 V/m, then

$$E_x = 0.1\cos(2\pi 10^6 t - .19z)\ \text{V/m}$$

$$H_y = \frac{E_x}{\eta} = 2.4\times 10^{-3}\cos(2\pi 10^6 t - .19z)\ \text{A/m}$$

✔ **D11.3.** A 9.375-GHz uniform plane wave is propagating in polyethylene (see Appendix C). If the amplitude of the electric field intensity is 500 V/m and the material is assumed to be lossless, find: (a) the phase constant; (b) the wavelength in the polyethylene; (c) the velocity of propagation; (d) the intrinsic impedance; (e) the amplitude of the magnetic field intensity.

Ans. 295 rad/m; 2.13 cm; 1.99×10^8 m/s; 251 Ω; 1.99 A/m

Example 11.4

We again consider plane wave propagation in water, but at the much higher microwave frequency of 2.5 GHz. At frequencies in this range and higher, dipole relaxation and resonance phenomena[3] in the water molecules become important. Real and imaginary parts of the permittivity are present, and both vary with frequency. At frequencies below that of visible light, the two mechanisms together produce an ϵ'' that increases with increasing frequency, reaching a local maximum in the vicinity of 10^{10} Hz. ϵ' decreases with increasing frequency. Ref. 3 provides specific details. At 2.5 GHz, dipole relaxation effects dominate. The permittivity values are $\epsilon'_R = 78$ and $\epsilon''_R = 7$. From(35), we have

$$\alpha = \frac{(2\pi\times 2.5\times10^9)\sqrt{78}}{(3.0\times10^8)\sqrt{2}}\left(\sqrt{1+\left(\frac{7}{78}\right)^2}-1\right)^{1/2} = 21\ \text{Np/m}$$

The first calculation demonstrates the operating principle of the *microwave oven*. Almost all foods contain water, and so can be cooked when incident microwave radiation is absorbed and converted into heat. Note that the field will attenuate to a value of e^{-1} times its initial value at a distance of $1/\alpha = 4.8$ cm. This distance is called the *penetration depth* of the material, and of course is frequency-dependent. The 4.8 cm depth is reasonable for cooking food, since it would lead to a temperature rise that is fairly uniform throughout the depth of the material. At much higher frequencies, where ϵ'' is larger, the penetration depth decreases, and too much power is absorbed at the surface; at lower frequencies, the penetration depth increases, and not enough overall absorption occurs. Commercial microwave ovens operate at frequencies in the vicinity of 2.5 GHz.

Using (36), in a calculation very similar to that for α, we find $\beta = 464$ rad/m. The wavelength is $\lambda = 2\pi/\beta = 1.4$ cm, whereas in free space this would have been $\lambda_0 = c/f = 12$ cm.

[3] These mechanisms and how they produce a complex permittivity are described in Appendix D. Additionally, the reader is referred to pp. 73–84 in Ref. 1 and pp. 678–682 in Ref. 2 for general treatments of relaxation and resonance effects on wave propagation. Discussions and data that are specific to water are presented in Ref. 3, pp. 314–316.

Using (39), the intrinsic impedance is found to be

$$\eta = \frac{377}{\sqrt{78}} \frac{1}{\sqrt{1 - j(7/78)}} = 43 + j1.9 = 43\angle 2.6^\circ \ \Omega$$

and E_x leads H_y in time by 2.6° at every point.

We next consider the case of conductive materials, in which currents are formed by the motion of free electrons or holes under the influence of an electric field. The governing relation is $\mathbf{J} = \sigma \mathbf{E}$, where σ is the material conductivity. With finite conductivity, the wave loses power through resistive heating of the material.We look for an interpretation of the complex permittivity as it relates to the conductivity. Consider the Maxwell curl equation (8) which, using (33), becomes:

$$\nabla \times \mathbf{H}_s = j\omega(\epsilon' - j\epsilon'')\mathbf{E}_s = \omega\epsilon''\mathbf{E}_s + j\omega\epsilon'\mathbf{E}_s \tag{44}$$

This equation can be expressed in a more familiar way, in which conduction current is included:

$$\nabla \times \mathbf{H}_s = \mathbf{J}_s + j\omega\epsilon\mathbf{E}_s \tag{45}$$

We next use $\mathbf{J}_s = \sigma\mathbf{E}_s$, and interpret ϵ in (41) as ϵ'. Eq. (45) then becomes:

$$\nabla \times \mathbf{H}_s = (\sigma + j\omega\epsilon')\mathbf{E}_s = \mathbf{J}_{\sigma s} + \mathbf{J}_{ds} \tag{46}$$

which we have expressed in terms of conduction current density, $\mathbf{J}_{\sigma s} = \sigma\mathbf{E}_s$, and displacement current density, $\mathbf{J}_{ds} = j\omega\epsilon'\mathbf{E}_s$. Comparing Eqs. (44) and (46), we find that in a conductive medium:

$$\boxed{\epsilon'' = \frac{\sigma}{\omega}} \tag{47}$$

Let us now turn our attention to the case of a dielectric material in which the loss is very small. The criterion by which we would judge whether or not the loss is small is the magnitude of the loss tangent, ϵ''/ϵ'. This parameter will have a direct influence on the attenuation coefficient, α, as seen from Eq. (35). In the case of conducting media in which (47) holds, the loss tangent becomes $\sigma/\omega\epsilon'$. By inspecting (46), we see that the ratio of condution current density to displacement current density magnitudes is

$$\frac{J_{\sigma s}}{J_{ds}} = \frac{\epsilon''}{j\epsilon'} = \frac{\sigma}{j\omega\epsilon'} \tag{48}$$

That is, these two vectors point in the same direction in space, but they are 90° out of phase in time. Displacement current density leads conduction current density by 90°, just as the current through a capacitor leads the current through a resistor in parallel with it by 90° in an ordinary electric current. This phase relationship is shown in Fig. 11.2. The angle θ (not to be confused with the polar

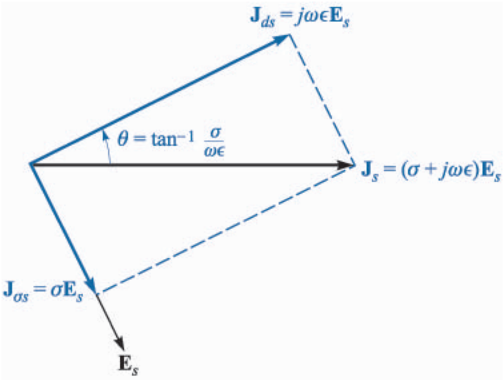


FIGURE 11.2
The time-phase relationship between $\mathbf{J}_{ds}$, $\mathbf{J}_{\sigma s}$, $\mathbf{J}_s$, and $\mathbf{E}_s$. The tangent of θ is equal to $\sigma/\omega\epsilon$, and $90^\circ - \theta$ is the common power-factor angle, or the angle by which $\mathbf{J}_s$ leads $\mathbf{E}_s$.

angle in spherical coordinates) may therefore be identified as the angle by which the displacement current density leads the total current density, and

$$\tan\theta = \frac{\epsilon''}{\epsilon'} = \frac{\sigma}{\omega\epsilon'} \tag{49}$$

The reasoning behind the term "loss tangent" is thus evident. Problem 16 at the end of the chapter indicates that the Q of a capacitor (its quality factor, not its charge) which incorporates a lossy dielectric is the reciprocal of the loss tangent.

If the loss tangent is small, then we may obtain useful approximations for the attenuation and phase constants, and the intrinsic impedance. Considering a conductive material, for which $\epsilon'' = \sigma/\omega$, (34) becomes

$$jk = j\omega\sqrt{\mu\epsilon'}\sqrt{1 - j\frac{\sigma}{\omega\epsilon'}} \tag{50}$$

We may expand the second radical using the binomial theorem

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{(2!)}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \ldots$$

where $|x| \ll 1$. We identify x as $-j\sigma/\omega\epsilon'$ and n as $1/2$, and thus

$$jk = j\omega\sqrt{\mu\epsilon'}\left[1 - j\frac{\sigma}{2\omega\epsilon'} + \frac{1}{8}\left(\frac{\sigma}{\omega\epsilon'}\right)^2 + \ldots\right] = \alpha + j\beta$$

Now

$$\alpha = \text{Re}(jk) \doteq j\omega\sqrt{\mu\epsilon'}\left(-j\frac{\sigma}{2\omega\epsilon'}\right) = \frac{\sigma}{2}\sqrt{\frac{\mu}{\epsilon'}} \tag{51}$$

and

$$\beta = \mathrm{Im}(jk) \doteq \omega\sqrt{\mu\epsilon'}\left[1 + \frac{1}{8}\left(\frac{\sigma}{\omega\epsilon'}\right)^2\right] \tag{52a}$$

or in many cases

$$\beta \doteq \omega\sqrt{\mu\epsilon'} \tag{52b}$$

Applying the binomial expansion to (39), we obtain

$$\eta \doteq \sqrt{\frac{\mu}{\epsilon'}}\left[1 - \frac{3}{8}\left(\frac{\sigma}{\omega\epsilon'}\right)^2 + j\frac{\sigma}{2\omega\epsilon'}\right] \tag{53a}$$

or

$$\eta \doteq \sqrt{\frac{\mu}{\epsilon'}}\left(1 + j\frac{\sigma}{2\omega\epsilon'}\right) \tag{53b}$$

The conditions under which the above approximations can be used depend on the desired accuracy, measured by how much the results deviate from those given by the exact formulas, (35) and (36). Deviations of no more than a few percent occur if $\sigma/\omega\epsilon' < 0.1$.

Example 11.5

As a comparison, we repeat the computations of Example 11.4, using the approximation formulas, (51), (52*b*), and (53*b*).

Solution. First, the loss tangent in this case is $\epsilon''/\epsilon' = 7/78 = 0.09$. Using (51), with $\epsilon'' = \sigma/\omega$, we have

$$\alpha \doteq \frac{\omega\epsilon''}{2}\sqrt{\frac{\mu}{\epsilon'}} = \frac{1}{2}(7 \times 8.85 \times 10^{12})(2\pi \times 2.5 \times 10^9)\frac{377}{\sqrt{78}} = 21\ \mathrm{cm}^{-1}$$

We then have, using (52*b*),

$$\beta \doteq (2\pi \times 2.5 \times 10^9)\sqrt{78}/(2.99 \times 10^8) = 464\ \mathrm{rad/m}$$

Finally, with (53*b*),

$$\eta \doteq \frac{377}{\sqrt{78}}\left(1 + j\frac{7}{2 \times 78}\right) = 43 + j1.9$$

These results are identical (within the accuracy limitations as determined by the given numbers) to those of Example 11.4. Small deviations will be found, as the reader can verify by repeating the calculations of both examples and expressing the results to four or five significant figures. As we know, this latter practice would not be meaningful since the given parameters were not specified with such accuracy. Such is often the case,

since measured values are not always known with high precision. Depending on how precise these values are, one can sometimes use a more relaxed judgement on when the approximation formulas can be used, by allowing loss tangent values that can be larger than 0.1 (but still less than 1).

✔ **D11.4.** Given a nonmagnetic material having $\epsilon'_R = 3.2$ and $\sigma = 1.5 \times 10^{-4}$ S/m, find numerical values at 3 MHz for the: (a) loss tangent; (b) attenuation constant; (c) phase constant; (d) intrinsic impedance.

Ans. 0.28; 0.016 Np/m; 0.11 rad/m; $207\angle 7.8^\circ\ \Omega$

✔ **D11.5.** Consider a material for which $\mu_R = 1$, $\epsilon'_R = 2.5$, and the loss tangent is 0.12. If these three values are constant with frequency in the range $0.5\text{ MHz} \le f \le 100$ MHz, calculate: (a) σ at 1 and 75 MHz; (b) λ at 1 and 75 MHz; (c) v_p at 1 and 75 MHz.

Ans. 1.67×10^{-5} and 1.25×10^{-3} S/m; 190 and 2.53 m; 1.90×10^{8} m/s twice

11.3 THE POYNTING VECTOR AND POWER CONSIDERATIONS

In order to find the power in a uniform plane wave, it is necessary to develop a theorem for the electromagnetic field known as the Poynting theorem. It was originally postulated in 1884 by an English physicist, John H. Poynting.

Let us begin with Maxwell's equation,

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}$$

and dot each side of the equation with **E**,

$$\mathbf{E} \cdot \nabla \times \mathbf{H} = \mathbf{E} \cdot \mathbf{J} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}$$

We now make use of the vector identity,

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\mathbf{E} \cdot \nabla \times \mathbf{H} + \mathbf{H} \cdot \nabla \times \mathbf{E}$$

which may be proved by expansion in rectangular coordinates. Thus

$$\mathbf{H} \cdot \nabla \times \mathbf{E} - \nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{J} \cdot \mathbf{E} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}$$

But

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

and therefore

$$-\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{J} \cdot \mathbf{E} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}$$

or

$$-\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{J} \cdot \mathbf{E} + \epsilon \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} + \mu \mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t}$$

However,

$$\epsilon \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} = \frac{\epsilon}{2} \frac{\partial E^2}{\partial t} = \frac{\partial}{\partial t}\left(\frac{\epsilon E^2}{2}\right)$$

and

$$\mu \mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t} = \frac{\partial}{\partial t}\left(\frac{\mu H^2}{2}\right)$$

Thus

$$-\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{J} \cdot \mathbf{E} + \frac{\partial}{\partial t}\left(\frac{\epsilon E^2}{2} + \frac{\mu H^2}{2}\right)$$

Finally, we integrate throughout a volume,

$$-\int_{vol} \nabla \cdot (\mathbf{E} \times \mathbf{H}) dv = \int_{vol} \mathbf{J} \cdot \mathbf{E}\, dv + \frac{\partial}{\partial t} \int_{vol} \left(\frac{\epsilon E^2}{2} + \frac{\mu H^2}{2}\right) dv$$

and apply the divergence theorem to obtain

$$-\oint_S (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{S} = \int_{vol} \mathbf{J} \cdot \mathbf{E} dv + \frac{\partial}{\partial t} \int_{vol} \left(\frac{\epsilon E^2}{2} + \frac{\mu H^2}{2}\right) dv \qquad (54)$$

If we assume that there are no sources within the volume, then the first integral on the right is the total (but instantaneous) ohmic power dissipated within the volume. If sources are present within the volume, then the result of integrating over the volume of the source will be positive if power is being delivered *to* the source, but it will be negative if power is being delivered *by* the source.

The integral in the second term on the right is the total energy stored in the electric and magnetic fields,[4] and the partial derivatives with respect to time cause this term to be the time rate of increase of energy stored within this volume, or the instantaneous power going to increase the stored energy within this volume. The sum of the expressions on the right must therefore be the total power flowing *into* this volume, and thus the total power flowing *out* of the volume is

$$\oint_S (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{S}$$

[4] This is the expression for magnetic field energy that we have been anticipating since Chap. 9.

where the integral is over the closed surface surrounding the volume. The cross product $\mathbf{E} \times \mathbf{H}$ is known as the Poynting vector, $\mathcal{P}$,

$$\boxed{\mathcal{P} = \mathbf{E} \times \mathbf{H}} \tag{55}$$

which is interpreted as an instantaneous power density, measured in watts per square meter ($\mathrm{W/m^2}$). This interpretation is subject to the same philosophical considerations as was the description of $(\mathbf{D} \cdot \mathbf{E})/2$ or $(\mathbf{B} \cdot \mathbf{H})/2$ as energy densities. We can show rigorously only that the integration of the Poynting vector over a closed surface yields the total power crossing the surface in an outward sense. This interpretation as a power density does not lead us astray, however, especially when applied to sinusoidally varying fields. Problem 11.18 indicates that strange results may be found when the Poynting vector is applied to time-constant fields.

The direction of the vector $\mathcal{P}$ indicates the direction of the instantaneous power flow at the point, and many of us think of the Poynting vector as a "pointing" vector. This homonym, while accidental, is correct.

Since $\mathcal{P}$ is given by the cross product of $\mathbf{E}$ and $\mathbf{H}$, the direction of power flow at any point is normal to both the $\mathbf{E}$ and $\mathbf{H}$ vectors. This certainly agrees with our experience with the uniform plane wave, for propagation in the $+z$ direction was associated with an E_x and H_y component,

$$E_x \mathbf{a_x} = H_y \mathbf{a_y} = \mathcal{P}_z \mathbf{a_z}$$

In a perfect dielectric, these $\mathbf{E}$ and $\mathbf{H}$ fields are given by

$$E_x = E_{x0} \cos(\omega t - \beta z)$$

$$H_y = \frac{E_{x0}}{\eta} \cos(\omega t - \beta z)$$

and thus

$$\mathcal{P}_z = \frac{E_{x0}^2}{\eta} \cos^2(\omega t - \beta z)$$

To find the time-average power density, we integrate over one cycle and divide by the period $T = 1/f$,

$$\begin{aligned}
\mathcal{P}_{z,av} &= \frac{1}{T} \int_0^T \frac{E_{x0}^2}{\eta} \cos^2(\omega t - \beta z)\, dt \\
&= \frac{1}{2T} \frac{E_{x0}^2}{\eta} \int_0^T [1 + \cos(2\omega t - 2\beta z)]\, dt \\
&= \frac{1}{2T} \frac{E_{x0}^2}{\eta} \left[t + \frac{1}{2\omega} \sin(2\omega t - 2\beta z) \right]_0^T
\end{aligned}$$

and

$$\mathcal{P}_{z,av} = \frac{1}{2}\frac{E_{x0}^2}{\eta} \text{ W/m}^2 \tag{56}$$

If we were using root-mean-square values instead of peak amplitudes, then the factor 1/2 would not be present.

Finally, the average power flowing through any area S normal to the z axis is[5]

$$P_{z,av} = \frac{1}{2}\frac{E_{x0}^2}{\eta} S \text{ W}$$

In the case of a lossy dielectric, E_x and H_y are not in time phase. We have

$$E_x = E_{x0}e^{-\alpha z}\cos(\omega t - \beta z)$$

If we let

$$\eta = |\eta|\angle\theta_\eta$$

then we may write the magnetic field intensity as

$$H_y = \frac{E_{x0}}{|\eta|}e^{-\alpha z}\cos(\omega t - \beta z - \theta_\eta)$$

Thus,

$$\mathcal{P}_z = E_x H_y = \frac{E_{x0}^2}{|\eta|}e^{-2\alpha z}\cos(\omega t - \beta z)\cos(\omega t - \beta z - \theta_\eta)$$

Now is the time to use the identity $\cos A \cos B \equiv 1/2\cos(A+B) + 1/2\cos(A-B)$, improving the form of the last equation considerably,

$$\mathcal{P}_z = \frac{1}{2}\frac{E_{x0}^2}{|\eta|}e^{-2\alpha z}[\cos(2\omega t - 2\beta z - 2\theta_\eta) + \cos\theta_\eta]$$

We find that the power density has only a second-harmonic component and a dc component. Since the first term has a zero average value over an integral number of periods, the time-average value of the Poynting vector is

$$\mathcal{P}_{z,av} = \frac{1}{2}\frac{E_{x0}^2}{|\eta|}e^{-2\alpha z}\cos\theta_\eta$$

Note that the power density attenuates as $e^{-2\alpha z}$, whereas E_x and H_y fall off as $e^{-\alpha z}$.

We may finally observe that the above expression for $\mathcal{P}_{z,av}$ can be obtained very easily by using the phasor forms of the electric and magnetic fields:

---

[5] We shall use P for power as well as for the polarization of the medium. If they both appear in the same equation in this book, it is an error.

$$\mathcal{P}_{z,av} = \frac{1}{2}\text{Re}(\mathbf{E_s} \times \mathbf{H_s^*}) \quad \text{W/m}^2 \tag{57}$$

where in the present case

$$\mathbf{E_s} = E_{x0}e^{-j\beta z}\mathbf{a_x}$$

and

$$\mathbf{H_s^*} = \frac{E_{x0}}{\eta^*}e^{+j\beta z}\mathbf{a_y} = \frac{E_{x0}}{|\eta|}e^{j\theta}e^{+j\beta z}\mathbf{a_y}$$

where E_{x0} has been assumed real. Eq. (57) applies to any sinusoidal electromagnetic wave, and gives both the magnitude and direction of the time-average power density.

✔ **D11.6.** At frequencies of 1, 100, and 3000 MHz, the dielectric constant of ice made from pure water has values of 4.15, 3.45, and 3.20, respectively, while the loss tangent is 0.12, 0.035, and 0.0009, also respectively. If a uniform plane wave with an amplitude of 100 V/m at $z = 0$ is propagating through such ice, find the time-average power density at $z = 0$ and $z = 10$ m for each frequency.

Ans. 27.1 and 25.7 W/m^2; 24.7 and 6.31 W/m^2; 23.7 and 8.63 W/m^2.

11.4 PROPAGATION IN GOOD CONDUCTORS: SKIN EFFECT

As an additional study of propagation with loss, we shall investigate the behavior of a good conductor when a uniform plane wave is established in it. Rather than thinking of a source embedded in a block of copper and launching a wave in that material, we should be more interested in a wave that is established by an electromagnetic field existing in some external dielectric that adjoins the conductor surface. We shall see that the primary transmission of energy must take place in the region *outside* the conductor, because all time-varying fields attenuate very quickly *within* a good conductor.

The good conductor has a high conductivity and large conduction currents. The energy represented by the wave traveling through the material therefore decreases as the wave propagates because ohmic losses are continuously present. When we discussed the loss tangent, we saw that the ratio of conduction current density to the displacement current density in a conducting material is given by $\sigma/\omega\epsilon'$. Choosing a poor metallic conductor and a very high frequency as a conservative example, this ratio[6] for nichrome ($\sigma \doteq 10^6$) at 100 MHz is about 2×10^8. Thus we have a situation where $\sigma/\omega\epsilon' \gg 1$, and we should be able to make several very good approximations to find α, β, and η for a good conductor.

[6] It is customary to take $\epsilon' = \epsilon_0$ for metallic conductors.

The general expression for the propagation constant is, from (50),

$$jk = j\omega\sqrt{\mu\epsilon'}\sqrt{1 - j\frac{\sigma}{\omega\epsilon'}}$$

which we immediately simplify to obtain

$$jk = j\omega\sqrt{\mu\epsilon'}\sqrt{-j\frac{\sigma}{\omega\epsilon'}}$$

or

$$jk = j\sqrt{-j\omega\mu\sigma}$$

But

$$-j = 1\angle -90^\circ$$

and

$$\sqrt{1\angle -90^\circ} = 1\angle -45^\circ = \frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}$$

Therefore

$$jk = j\left(\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}\right)\sqrt{\omega\mu\sigma}$$

or

$$jk = (j1 + 1)\sqrt{\pi f\mu\sigma} = \alpha + j\beta \tag{58}$$

Hence

$$\boxed{\alpha = \beta = \sqrt{\pi f\mu\sigma}} \tag{59}$$

Regardless of the parameters μ and σ of the conductor or of the frequency of the applied field, α and β are equal. If we again assume only an E_x component traveling in the $+z$ direction, then

$$E_x = E_{x0}e^{-z\sqrt{\pi f\mu\sigma}}\cos(\omega t - z\sqrt{\pi f\mu\sigma}) \tag{60}$$

We may tie this field in the conductor to an external field at the conductor surface. We let the region $z > 0$ be the good conductor and the region $z < 0$ be a perfect dielectric. At the boundary surface $z = 0$, (60) becomes

$$E_x = E_{x0}\cos\omega t \qquad (z = 0)$$

This we shall consider as the source field that establishes the fields within the conductor. Since displacement current is negligible,

$$\mathbf{J} = \sigma\mathbf{E}$$

Thus, the conduction current density at any point within the conductor is directly related to **E**:

$$J_x = \sigma E_x = \sigma E_{x0} e^{-z\sqrt{\pi f \mu \sigma}} \cos(\omega t - z\sqrt{\pi f \mu \sigma}) \qquad (61)$$

Equations (60) and (61) contain a wealth of information. Considering first the negative exponential term, we find an exponential decrease in the conduction current density and electric field intensity with penetration into the conductor (away from the source). The exponential factor is unity at $z = 0$ and decreases to $e^{-1} = 0.368$ when

$$z = \frac{1}{\sqrt{\pi f \mu \sigma}}$$

This distance is denoted by δ and is termed the *depth of penetration*, or the *skin depth*,

$$\boxed{\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} = \frac{1}{\alpha} = \frac{1}{\beta}} \qquad (62)$$

It is an important parameter in describing conductor behavior in electromagnetic fields. To get some idea of the magnitude of the skin depth, let us consider copper, $\sigma = 5.8 \times 10^7$ S/m, at several different frequencies. We have

$$\delta_{Cu} = \frac{0.066}{\sqrt{f}}$$

At a power frequency of 60 Hz, $\delta_{Cu} = 8.53$ mm, or about 1/3 in. Remembering that the power density carries an exponential term $e^{-2\alpha z}$, we see that the power density is multiplied by a factor of $0.368^2 = 0.135$ for every 8.53 mm of distance into the copper.

At a microwave frequency of 10,000 MHz, δ is 6.61×10^{-4} mm. Stated more generally, all fields in a good conductor such as copper are essentially zero at distances greater than a few skin depths from the surface. Any current density or electric field intensity established at the surface of a good conductor decays rapidly as we progress into the conductor. Electromagnetic energy is not transmitted in the interior of a conductor; it travels in the region surrounding the conductor, while the conductor merely guides the waves. We shall consider guided propagation in more detail in Chapters 13 and 14.

Suppose we have a copper bus bar in the substation of an electric utility company which we wish to have carry large currents, and we therefore select dimensions of 2 by 4 in. Then much of the copper is wasted, for the fields are greatly reduced in one skin depth, about 1/3 in.[7] A hollow conductor with a wall

[7] This utility company operates at 60 Hz.

thickness of about 1/2 in would be a much better design. Although we are applying the results of an analysis for an infinite planar conductor to one of finite dimensions, the fields are attenuated in the finite-size conductor in a similar (but not identical) fashion.

The extremely short skin depth at microwave frequencies shows that only the surface coating of the guiding conductor is important. A piece of glass with an evaporated silver surface 0.0001 in thick is an excellent conductor at these frequencies.

Next, let us determine expressions for the velocity and wavelength within a good conductor. From (62), we already have

$$\alpha = \beta = \frac{1}{\delta} = \sqrt{\pi f \mu \sigma}$$

Then, since

$$\beta = \frac{2\pi}{\lambda}$$

we find the wavelength to be

$$\boxed{\lambda = 2\pi\delta} \tag{63}$$

Also, recalling that

$$v_p = \frac{\omega}{\beta}$$

we have

$$\boxed{v_p = \omega\delta} \tag{64}$$

For copper at 60 Hz, $\lambda = 5.36$ cm and $v_p = 3.22$ m/s, or about 7.2 mi/h. A lot of us can run faster than that. In free space, of course, a 60-Hz wave has a wavelength of 3100 mi and travels at the velocity of light.

Example 11.6

Let us again consider wave propagation in water, but this time we will consider seawater. The primary difference between seawater and fresh water is of course the salt content. Sodium chloride dissociates in water to form Na^+ and Cl^- ions, which, being charged, will move when forced by an electric field. Seawater is thus conductive, and so will attenuate electromagnetic waves by this mechanism. At frequencies in the vicinity of 10^7 Hz and below, the bound charge effects in water discussed earlier are negligible, and losses in seawater arise principally from the salt-associated conductivity. We consider an incident wave of frequency 1 MHz. We wish to find the skin depth, wavelength, and phase velocity. In seawater, $\sigma = 4$ S/m, and $\epsilon'_R = 81$.

Solution. We first evaluate the loss tangent, using the given data:

$$\frac{\sigma}{\omega\epsilon'} = \frac{4}{(2\pi \times 10^6)(81)(8.85 \times 10^{-12})} = 8.9 \times 10^2 \gg 1$$

Thus, seawater is a good conductor at 1 MHz (and at frequencies lower than this). The skin depth is

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} = \frac{1}{\sqrt{(\pi \times 10^6)(4\pi \times 10^{-7})(4)}} = 0.25 \text{ m} = 25 \text{ cm}$$

Now

$$\lambda = 2\pi\delta = 1.6 \text{ m}$$

and

$$v_p = \omega\delta = (2\pi \times 10^6)(0.25) = 1.6 \times 10^6 \text{ m/sec}$$

In free space, these values would have been $\lambda = 300$ m and of course $v = c$.

With a 25 cm skin depth, it is obvious that radio frequency communication in seawater is quite impractical. Notice however that δ varies as $1/\sqrt{f}$, so that things will improve at lower frequencies. For example, if we use a frequency of 10 Hz in the extremely low frequency (ELF) range, the skin depth is increased over that at 1 MHz by a factor of $\sqrt{10^6/10}$, so that

$$\delta(10Hz) \doteq 80 \text{ m}$$

The corresponding wavelength is $\lambda = 2\pi\delta \doteq 500$ m. Frequencies in the ELF range are in fact used for submarine communications, chiefly between gigantic ground-based antennas (required since the free-space wavelength associated with 10 Hz is 3×10^7 m) and submarines, from which a suspended wire antenna of length shorter than 500 m is sufficient to receive the signal. The drawback is that signal data rates at ELF are so slow that a single word can take several minutes to transmit. Typically, ELF signals are used to tell the submarine to implement emergency procedures, or to come near the surface in order to receive a more detailed message via satellite.

We next turn our attention to finding the magnetic field, H_y, associated with E_x. To do so, we need an expression for the intrinsic impedance of a good conductor. We begin with Eq. (39), Sec. 11.2, with $\epsilon'' = \sigma/\omega$,

$$\eta = \sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon'}}$$

Since $\sigma \gg \omega\epsilon'$, we have

$$\eta = \sqrt{\frac{j\omega\mu}{\sigma}}$$

which may be written as

$$\boxed{\eta = \frac{\sqrt{2}\angle 45^\circ}{\sigma\delta} = \frac{1}{\sigma\delta} + j\frac{1}{\sigma\delta}} \qquad (65)$$

Thus, if we write (60) in terms of the skin depth,

$$E_x = E_{x0} e^{-z/\delta} \cos\left(\omega t - \frac{z}{\delta}\right) \tag{66}$$

then

$$H_y = \frac{\sigma \delta E_{x0}}{\sqrt{2}} e^{-z/\delta} \cos\left(\omega t - \frac{z}{\delta} - \frac{\pi}{4}\right) \tag{67}$$

and we see that the maximum amplitude of the magnetic field intensity occurs one-eighth of a cycle later than the maximum amplitude of the electric field intensity at every point.

From (66) and (67) we may obtain the time-average Poynting vector by applying (57),

$$\mathcal{P}_{z,av} = \frac{1}{2} \frac{\sigma \delta E_{x0}^2}{\sqrt{2}} e^{-2z/\delta} \cos\left(\frac{\pi}{4}\right)$$

or

$$\boxed{\mathcal{P}_{z,av} = \frac{1}{4} \sigma \delta E_{x0}^2 e^{-2z/\delta}}$$

We again note that in a distance of one skin depth the power density is only $e^{-2} = 0.135$ of its value at the surface.

The total power loss in a width $0 < y < b$ and length $0 < x < L$ in the direction of the current, as shown in Fig. 11.3, is obtained by finding the power crossing the conductor surface within this area,

$$P_{L,av} = \int_S \mathcal{P}_{z,av} dS = \int_0^b \int_0^L \frac{1}{4} \sigma \delta E_{x0}^2 e^{-2z/\delta}|_{z=0} dx\, dy = \frac{1}{4} \sigma \delta b L E_{x0}^2$$

In terms of the current density J_{x0} at the surface,

$$J_{x0} = \sigma E_{x0}$$

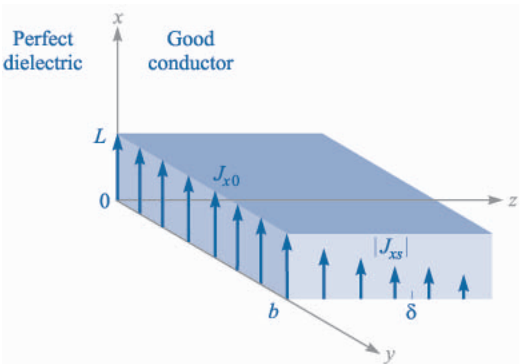


FIGURE 11.3
The current density $J_x = J_{x0} e^{-z/\delta} e^{-jz/\delta}$ decreases in magnitude as the wave propagates into the conductor. The average power loss in the region $0 < x < L$, $0 < y < b$, $z > 0$, is $\delta b L J_{x0}^2 / 4\sigma$ watts.

we have

$$P_{L,av} = \frac{1}{4\sigma}\delta b L J_{x0}^2 \tag{68}$$

Now let us see what power loss would result if the *total* current in a width b were distributed *uniformly* in one skin depth. To find the total current, we integrate the current density over the infinite depth of the conductor,

$$I = \int_0^\infty \int_0^b J_x dy\, dz$$

where

$$J_x = J_{x0} e^{-z/\delta} \cos\left(\omega t - \frac{z}{\delta}\right)$$

or in complex exponential notation to simplify the integration,

$$\begin{aligned} J_{xs} &= J_{x0} e^{-z/\delta} e^{-jz/\delta} \\ &= J_{x0} e^{-(1+j1)z/\delta} \end{aligned}$$

Therefore,

$$\begin{aligned} I_s &= \int_0^\infty \int_0^b J_{x0} e^{-(1+j1)z/\delta} dy\, dz \\ &= J_{x0} b e^{-(1+j1)z/\delta} \left.\frac{-\delta}{1+j1}\right|_0^\infty \\ &= \frac{J_{x0} b\delta}{1+j1} \end{aligned}$$

and

$$I = \frac{J_{x0} b\delta}{\sqrt{2}} \cos\left(\omega t - \frac{\pi}{4}\right)$$

If this current is distributed with a uniform density J' throughout the cross section $0 < y < b$, $0 < z < \delta$, then

$$J' = \frac{J_{x0}}{\sqrt{2}} \cos\left(\omega t - \frac{\pi}{4}\right)$$

The ohmic power loss per unit volume is $\mathbf{J} \cdot \mathbf{E}$, and thus the total instantaneous power dissipated in the volume under consideration is

$$P_L = \frac{1}{\sigma}(J')^2 bL\delta = \frac{J_{x0}^2}{2\sigma} bL\delta \cos^2\left(\omega t - \frac{\pi}{4}\right)$$

The time-average power loss is easily obtained, since the average value of the cosine-squared factor is one-half,

$$P_{L,av} = \frac{1}{4\sigma} J_{x0}^2 bL\delta \tag{69}$$

Comparing (68) and (69), we see that they are identical. Thus the average power loss in a conductor with skin effect present may be calculated by assuming that the total current is distributed uniformly in one skin depth. In terms of resistance, we may say that the resistance of a width b and length L of an infinitely thick slab with skin effect is the same as the resistance of a rectangular slab of width b, length L, and thickness δ without skin effect, or with uniform current distribution.

We may apply this to a conductor of circular cross section with little error, provided that the radius a is much greater than the skin depth. The resistance at a high frequency where there is a well-developed skin effect is therefore found by considering a slab of width equal to the circumference $2\pi a$ and thickness δ. Hence

$$R = \frac{L}{\sigma S} = \frac{L}{2\pi a\sigma\delta} \tag{70}$$

A round copper wire of 1 mm radius and 1 km length has a resistance at direct current of

$$R_{dc} = \frac{10^3}{\pi 10^{-6}(5.8 \times 10^7)} = 5.48\ \Omega$$

At 1 MHz, the skin depth is 0.066 mm. Thus $\delta \ll a$, and the resistance at 1 MHz is found by (70),

$$R = \frac{10^3}{2\pi 10^{-3}(5.8 \times 10^7)(0.066 \times 10^{-3})} = 41.5\ \Omega$$

✔ **D11.7.** A steel pipe is constructed of a material for which $\mu_R = 180$ and $\sigma = 4 \times 10^6$ S/m. The two radii are 5 and 7 mm, and the length is 75 m. If the total current $I(t)$ carried by the pipe is $8\cos\omega t$ A, where $\omega = 1200\pi$ rad/s, find: (a) the skin depth; (b) the effective resistance; (c) the dc resistance; (d) the time-average power loss.

Ans. 0.766 mm; 0.557 Ω; 0.249 Ω; 17.82 W

11.5 WAVE POLARIZATION

In the previous sections, we have treated uniform plane waves in which the electric and magnetic field vectors were assumed to lie in fixed directions. Specifically, with the wave propagating along the z axis, **E** was taken to lie along x, which then required **H** to lie along y. This orthogonal relationship between **E**, **H**, and $\mathcal{P}$ is always true for a uniform plane wave. The directions of **E** and **H** within the plane perpendicular to $\mathbf{a_z}$ may change, however, as functions of time and position, depending on how the wave was generated, or on what type of medium it is propagating through. Thus a complete description of

an electromagnetic wave would not only include parameters such as its wavelength, phase velocity, and power, but also a statement of the instantaneous orientation of its field vectors. *The wave polarization is defined as its electric field vector orientation as a function of time, at a fixed position in space.* Specifying only the electric field direction is sufficient, since magnetic field is readily found from **E** using Maxwell's equations.

In the waves we have previously studied, **E** was in a fixed straight orientation for all times and positions. Such a wave is said to be *linearly polarized*. We have taken **E** to lie along the x axis, but the field could be oriented in any fixed direction in the $x-y$ plane and be linearly polarized. For positive z-propagation, the wave would generally have its electric field phasor expressed as

$$\mathbf{E_s} = (E_{x0}\mathbf{a_x} + E_{y0}\mathbf{a_y})e^{-\alpha z}e^{-j\beta z} \tag{71}$$

where E_{x0} and E_{y0} are constant amplitudes along x and y. The associated magnetic field is readily found by determining its x and y components directly from those of $\mathbf{E_s}$. Specifically, $\mathbf{H_s}$ for the wave of Eq. (71) is

$$\mathbf{H_s} = [H_{x0}\mathbf{a_x} + H_{y0}\mathbf{a_y}]e^{-\alpha z}e^{-j\beta z} = \left[-\frac{E_{y0}}{\eta}\mathbf{a_x} + \frac{E_{x0}}{\eta}\mathbf{a_y}\right]e^{-\alpha z}e^{-j\beta z} \tag{72}$$

The two fields are sketched in Fig. 11.4. The figure demonstrates the reason for the minus sign in the term involving E_{y0} in Eq. (72). The direction of power flow, given by $\mathbf{E} \times \mathbf{H}$, is in the positive z direction in this case. A component of **E** in the positive y direction would require a component of **H** in the negative x direction—thus the minus sign. Using (71) and (72), the power density in the wave is found using (57):

$$\begin{aligned}\mathcal{P}_{z,av} &= \frac{1}{2}\text{Re}\{\mathbf{E_s} \times \mathbf{H_s^*}\} = \frac{1}{2}\text{Re}\{E_{x0}H_{y0}^*(\mathbf{a_x} \times \mathbf{a_y}) + E_{y0}H_{x0}^*(\mathbf{a_y} \times \mathbf{a_x})\}e^{-2\alpha z} \\ &= \frac{1}{2}\text{Re}\left\{\frac{E_{x0}E_{x0}^*}{\eta^*} + \frac{E_{y0}E_{y0}^*}{\eta^*}\right\}e^{-2\alpha z}\mathbf{a_z} = \frac{1}{2}\text{Re}\left\{\frac{1}{\eta^*}\right\}(|E_{x0}|^2 + |E_{y0}|^2)e^{-2\alpha z}\mathbf{a_z}\ \text{W/m}^2\end{aligned}$$

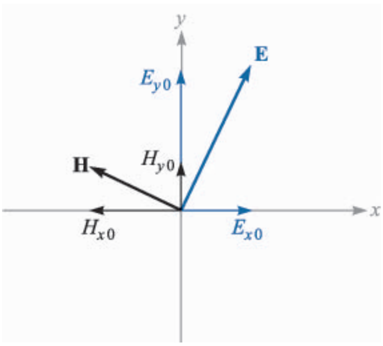


FIGURE 11.4
Electric and magnetic field configuration for a general linearly polarized plane wave, propagating in the forward z direction (out of the page). Field components correspond to those in Eqs. (71) and (72).

This result demonstrates the idea that our linearly polarized plane wave can be considered as two distinct plane waves having x and y polarizations, and whose electric fields are adding *in phase* to produce the total **E**. The same is true for the magnetic field components. This is a critical point in understanding wave polarization, in that *any polarization state can be described in terms of mutually perpendicular components of the electric field and their relative phasing*.

We next consider the effect of a phase difference, ϕ, between E_{x0} and E_{y0}, where $\phi < \pi/2$. For simplicity, we will consider propagation in a lossless medium. The total field in phasor form is

$$\mathbf{E_s} = (E_{x0}\mathbf{a_x} + E_{y0}e^{j\phi}\mathbf{a_y})e^{-j\beta z} \tag{73}$$

Again, to aid in visualization, we convert this wave to real instantaneous form by multiplying by $e^{j\omega t}$ and taking the real part:

$$\mathbf{E}(z, t) = E_{x0}\cos(\omega t - \beta z)\mathbf{a_x} + E_{y0}\cos(\omega t - \beta z + \phi)\mathbf{a_y} \tag{74}$$

where we have assumed that E_{x0} and E_{y0} are real. Suppose we set $t = 0$, in which case (74) becomes (using $\cos(-x) = \cos(x)$)

$$\mathbf{E}(z, 0) = E_{x0}\cos(\beta z)\mathbf{a_x} + E_{y0}\cos(\beta z - \phi)\mathbf{a_y} \tag{75}$$

The component magnitudes of $\mathbf{E}(z, 0)$ are plotted as functions of z in Fig. 11.5. Since time is fixed at zero, the wave is frozen in position. An observer can move along the z axis, measuring the component magnitudes and thus the orientation of the total electric field at each point. Let's consider a crest of E_x, indicated as point a in the figure. If ϕ were zero, E_y would have a crest at the same location. Since ϕ is not zero (and positive), the crest of E_y that would otherwise occur at point a is now displaced to point b further down z. The two points are separated by distance ϕ/β. E_y thus *lags behind* E_x when considering the *spatial* dimension.

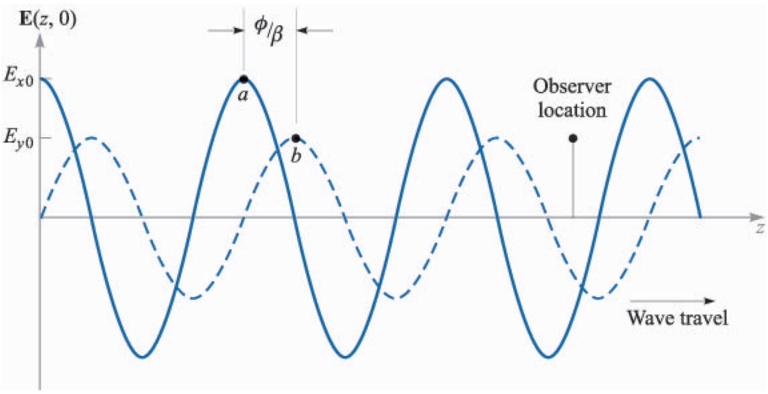


FIGURE 11.5
Plots of the electric field component magnitudes in Eq. (75) as functions of z. Note that the y component lags the x component in z. As time increases, both waves travel to the right, as per Eq. (74). Thus to an observer at a fixed location, the y component leads in time.

Now suppose the observer stops at some location on the z axis, and time is allowed to move forward. Both fields now move in the positive z direction, as (74) indicates. But point b reaches the observer first, followed by point a. So we see that E_y *leads* E_x when considering the *time* dimension. In either case (fixed t and varying z, or vice-versa) the observer notes that the net field rotates about the z axis while its magnitude changes. Considering a starting point in z and t, at which the field has a given orientation and magnitude, the wave will return to the same orientation and magnitude at a distance of one wavelength in z (for fixed t) or at a time $t = 2\pi/\omega$ later (at a fixed z).

For illustration purposes, if we take the length of the field vector as a measure of its magnitude, we would find that at a fixed position, the tip of the vector would trace out the shape of an ellipse over time $t = 2\pi/\omega$. The wave is thus said to be *elliptically polarized*. Elliptical polarization is in fact the most general polarization state of a wave, since it encompasses any magnitude and phase difference between E_x and E_y. Linear polarization is a special case of elliptical polarization, in which the phase difference is zero.

Another special case of elliptical polarization occurs when $E_{x0} = E_{y0} = E_0$ and when $\phi = \pm\pi/2$. The wave in this case exhibits *circular polarization*. To see this, we incorporate the above restrictions into Eq. (74) to obtain

$$\begin{aligned} \mathbf{E}(z, t) &= E_0[\cos(\omega t - \beta z)\mathbf{a_x} + \cos(\omega t - \beta z \pm \pi/2)\mathbf{a_y}] \\ &= E_0[\cos(\omega t - \beta z)\mathbf{a_x} \mp \sin(\omega t - \beta z)\mathbf{a_y}] \end{aligned} \tag{76}$$

If we consider a fixed position along z (such as $z = 0$) and allow time to vary, (76), with $\phi = +\pi/2$, becomes

$$\mathbf{E}(0, t) = E_0[\cos(\omega t)\mathbf{a_x} - \sin(\omega t)\mathbf{a_y}] \tag{77}$$

If we choose $\phi = -\pi/2$ in (76), we obtain

$$\mathbf{E}(0, t) = E_0[\cos(\omega t)\mathbf{a_x} + \sin(\omega t)\mathbf{a_y}] \tag{78}$$

The field vector of Eq. (78) rotates in the counter-clockwise direction in the x, y plane, while maintaining constant amplitude E_0, and so the tip of the vector traces out a circle. Fig. 11.6 shows this behavior. Choosing $\phi = +\pi/2$ leads to (77), whose field vector rotates in the clockwise direction. The *handedness* of the circular polarization is associated with the rotation and propagation directions in the following manner: The wave exhibits *left circular polarization* (l.c.p.) if when orienting the left hand with the thumb in the direction of propagation, the fingers curl in the rotation direction of the field with time. The wave exhibits *right circular polarization* (r.c.p.) if with the right hand thumb in the propagation direction, the fingers curl in the field rotation direction. Thus, with forward z propagation, (77) describes a left circularly polarized wave, and (78) a right circularly polarized wave. The same convention is applied to elliptical polarization, in which the descriptions *left elliptical polarization* and *right elliptical polarization* are used.

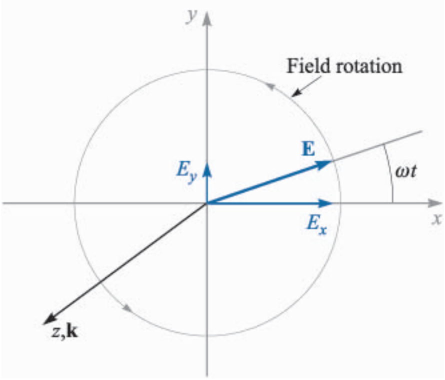


FIGURE 11.6
Electric field in the x, y plane of a right circularly polarized plane wave, as described by Eq. (78). As the wave propagates in the forward z direction, the field vector rotates counterclockwise in the x, y plane.

Using (76), the instantaneous angle of the field from the x direction can be found for any position along z through

$$\theta(z, t) = \tan^{-1}\left(\frac{E_y}{E_x}\right) = \tan^{-1}\left(\frac{\mp\sin(\omega t - \beta z)}{\cos(\omega t - \beta z)}\right) = \mp(\omega t - \beta z) \qquad (79)$$

where again the minus sign (yielding l.c.p. for positive z travel) applies for the choice of $\phi = +\pi/2$ in (76); the plus sign (yielding r.c.p. for positive z travel) is used if $\phi = -\pi/2$. If we choose $z = 0$, the angle becomes simply ωt, which reaches 2π (one complete rotation) at time $t = 2\pi/\omega$. If we chose $t = 0$ and allow z to vary, we form a corkscrew-like field pattern. One way to visualize this is to consider a spiral staircase-shaped pattern, in which the field lines (stair-steps) are perpendicular to the z (or staircase) axis. The relationship between this spatial field pattern and the resulting time behavior at fixed z as the wave propagates is shown in an artist's conception on the front cover.[8] Changing handedness of the polarization is accomplished by reversing the pitch of the corkscrew pattern. The spiral staircase model is only a visualization aid. It must be remembered that the wave is still a uniform plane wave, whose fields at any position along z are infinite in extent over the transverse plane.

There are many uses of circularly polarized waves. Perhaps the most obvious advantage is that reception of a wave having circular polarization does not depend on the antenna orientation in the plane normal to the propagation direction. Dipole antennas, for example, are required to be oriented along the electric field direction of the signal they receive. If circularly polarized signals are transmitted, the receiver orientation requirements are relaxed considerably. In optics, circularly polarized light can be passed through a polarizer of any orientation, thus yielding linearly polarized light in any direction (although one loses half the original power this way). Other uses involve treating linearly

[8] The artistic license enabling the cover wave to cast a shadow of itself may cause one to stop and think for a moment (but only for a moment).

polarized light as a superposition of circularly polarized waves, to be described below.

Circularly polarized light can be generated using an *anisotropic* medium—a material whose permittivity is a function of electric field direction. Many crystals have this property. A crystal orientation can be found such that along one direction (say, the x axis), the permittivity is lowest, while along the orthogonal direction (y axis), the permittivity is highest. The strategy is to input a linearly polarized wave with its field vector at 45 degrees to the x and y axes of the crystal. It will thus have equal-amplitude x and y components in the crystal, and these will now propagate in the z direction at different speeds. A phase difference (or *retardation*) accumulates between the components as they propagate, which can reach $\pi/2$ if the crystal is long enough. The wave at the output thus becomes circularly polarized. Such a crystal, cut to the right length, and used in this manner, is called a *quarter-wave plate*, since it introduces a relative phase shift of $\pi/2$ between E_x and E_y, which is equivalent to $\lambda/4$.

It is useful to express circularly polarized waves in phasor form. To do this, we note that (76) can be expressed as

$$\mathbf{E}(z,t) = \text{Re}\left\{E_0 e^{j\omega t} e^{-j\beta z}\left[\mathbf{a_x} + e^{\pm j\pi/2}\mathbf{a_y}\right]\right\}$$

Using the fact that $e^{\pm j\pi/2} = \pm j$, we identify the phasor form as:

$$\boxed{\mathbf{E}_s = E_0(\mathbf{a_x} \pm j\mathbf{a_y})e^{-j\beta z}} \qquad (80)$$

where the plus sign is used for left circular polarization, and the minus sign for right circular polarization. If the wave propagates in the negative z direction, we would have

$$\boxed{\mathbf{E}_s = E_0(\mathbf{a_x} \pm j\mathbf{a_y})e^{+j\beta z}} \qquad (81)$$

where in this case the positive sign applies to right circular polarization, and the minus sign to left circular polarization. The student is encouraged to verify this.

Example 11.7

Let us consider the result of superimposing left and right circularly polarized fields of the same amplitude, frequency, and propagation direction, but where a phase shift of δ radians exists between the two.

Solution. Taking the waves to propagate in the $+z$ direction, and introducing a relative phase, δ, the total phasor field is found, using (80):

$$\mathbf{E}_s^T = \mathbf{E}_s^R + \mathbf{E}_s^L = E_0[\mathbf{a_x} - j\mathbf{a_y}]e^{-j\beta z} + E_0[\mathbf{a_x} + j\mathbf{a_y}]e^{-j\beta z}e^{j\delta}$$

Grouping components together, this becomes

$$\mathbf{E}_s^T = E_0[(1 + e^{j\delta})\mathbf{a_x} - j(1 - e^{j\delta})\mathbf{a_y}]e^{-j\beta z}$$

Factoring out an overall phase term, $e^{j\delta/2}$, we obtain

$$\mathbf{E}_s^T = E_0 e^{j\delta/2}\left[(e^{-j\delta/2} + e^{j\delta/2})\mathbf{a_x} - j(e^{-j\delta/2} - e^{j\delta/2})\mathbf{a_y}\right]e^{-j\beta z}$$

From Euler's identity, we find that $e^{j\delta/2} + e^{-j\delta/2} = 2\cos\delta/2$, and $e^{j\delta/2} - e^{-j\delta/2} = 2j\sin\delta/2$. Using these relations, we obtain

$$\mathbf{E}_s^T = 2E_0[\cos(\delta/2)\mathbf{a_x} + \sin(\delta/2)\mathbf{a_y}]e^{-j(\beta z - \delta/2)} \tag{82}$$

We recognize (82) as the electric field of a *linearly polarized* wave, whose field vector is oriented at angle $\delta/2$ from the x axis.

The above example shows that any linearly polarized wave can be expressed as the sum of two circularly polarized waves of opposite handedness, and where the linear polarization direction is determined by the relative phase difference between the two waves. Such a representation is convenient (and necessary) when considering, for example, the propagation of linearly polarized light through media which contain organic molecules. These often exhibit spiral structures having left- or right-handed pitch, and will thus interact differently with left- or right-hand circular polarization. As a result, the left circular component could propagate at a different speed than the right circular component, and so the two waves will accumulate a phase difference as they propagate. The direction of the linearly polarized field vector at the output of the material will thus differ from the direction that it had at the input. The extent of this rotation can be used as a measurement tool to aid in material studies.

Polarization issues will become extremely important when we consider wave reflection in the next chapter.

SUGGESTED REFERENCES

1. Balanis, C. A.: "Advanced Engineering Electromagnetics," John Wiley and Sons, New York, 1989.
2. Ramo, S., J. R. Whinnery, and T. Van Duzer: "Fields and Waves in Communication Electronics," 3rd ed., John Wiley and Sons, New York, 1994.
3. Jackson, J. D.: "Classical Electrodynamics," 3rd ed., John Wiley and Sons, New York, 1999.
4. International Telephone and Telegraph Co., Inc.: "Reference Data for Radio Engineers," 7th ed., Howard W. Sams & Co., Indianapolis, Ind., 1985. This handbook has some excellent data on the properties of dielectric and insulating materials.

PROBLEMS

11.1 Show that $E_{xs} = Ae^{j(k_0 z+\phi)}$ is a solution of the vector Helmholtz equation, Sec. 11.1, Eq. (16), for $k_0 = \omega\sqrt{\mu_0\epsilon_0}$ and any ϕ and A.

11.2 Let $E(z, t) = 200 \sin 0.2z \cos 10^8 t\mathbf{a_x} + 500 \cos(0.2z + 50^\circ) \sin 10^8 t\mathbf{a_y}$ V/m. Find: (a) $\mathbf{E}$ at $P(0, 2, 0.6)$ at $t = 25$ ns; (b) $|E|$ at P at $t = 20$ ns; (c) E_s; (d) E_s at P.

11.3 An $\mathbf{H}$ field in free space is given as $\mathbf{H}(x, t) = 10 \cos(10^8 t - \beta x)\mathbf{a_y}$ A/m. Find: (a) β; (b) λ; (c) $\mathbf{E}(x, t)$ at $P(0.1, 0.2, 0.3)$ at $t = 1$ ns.

11.4 In phasor form, the electric field intensity of a uniform plane wave in free space is expressed as $\mathbf{E_s} = (40 - j30)e^{-j20z}\mathbf{a_x}$ V/m. Find: (a) ω; (b) β; (c) f; (d) λ; (e) $\mathbf{H_s}$; (f) $\mathbf{H}(z, t)$ at $P(6, -1, 0.07)$, $t = 71$ ps.

11.5 A 150-MHz uniform plane wave in free space is described by $\mathbf{H_s} = (4 + j10)(2\mathbf{a_x} + j\mathbf{a_y})e^{-j\beta z}$ A/m. (a) Find numerical values for ω, λ, and β. (b) Find $\mathbf{H}(z, t)$ at $t = 1.5$ ns, $z = 20$ cm. (c) What is $|E|_{max}$?

11.6 Let $\mu_R = \epsilon_R = 1$ for the field $\mathbf{E}(z, t) = (25\mathbf{a_x} - 30\mathbf{a_y}) \cos(\omega t - 50z)$ V/m. (a) Find ω. (b) Determine the displacement current density, $\mathbf{J_d}(z, t)$. (c) Find the total magnetic flux Φ passing through the rectangle defined by $0 < x < 1$, $y = 0$, $0 < z < 1$, at $t = 0$.

11.7 The phasor magnetic field intensity for a 400-MHz uniform plane wave propagating in a certain lossless material is $(2\mathbf{a_y} - j5\mathbf{a_z})e^{-j25x}$ A/m. Knowing that the maximum amplitude of $\mathbf{E}$ is 1500 V/m, find β, η, λ, v_p, ϵ_R, μ_R, and $\mathbf{H}(x, y, z, t)$.

11.8 Let the fields, $\mathbf{E}(z, t) = 1800 \cos(10^7 \pi t - \beta z)\mathbf{a_x}$ V/m and $\mathbf{H}(z, t) = 3.8 \cos(10^7 \pi t - \beta z)\mathbf{a_y}$ A/m, represent a uniform plane wave propagating at a velocity of 1.4×10^8 m/s in a perfect dielectric. Find: (a) β; (b) λ; (c) η; (d) μ_R; (e) ϵ_R.

11.9 A certain lossless material has $\mu_R = 4$ and $\epsilon_R = 9$. A 10-MHz uniform plane wave is propagating in the $\mathbf{a_y}$ direction with $E_{x0} = 400$ V/m and $E_{y0} = E_{z0} = 0$ at $P(0.6, 0.6, 0.6)$ at $t = 60$ ns. (a) Find β, λ, v_p, and η. (b) Find $E(t)$. (c) Find $H(t)$.

11.10 Given a 20-MHz uniform plane wave with $\mathbf{H_s} = (6\mathbf{a_x} - j2\mathbf{a_y})e^{-jz}$ A/m, assume propagation in a lossless medium characterized by $\epsilon_R = 5$ and an unknown μ_R. (a) Find λ, v_p, μ_R, and η. (b) Determine $\mathbf{E}$ at the origin at $t = 20$ ns.

11.11 A 2-GHz uniform plane wave has an amplitude $E_{y0} = 1.4$ kV/m at $(0, 0, 0, t = 0)$ and is propagating in the $\mathbf{a_z}$ direction in a medium where $\epsilon'' = 1.6 \times 10^{-11}$ F/m, $\epsilon' = 3.0 \times 10^{-11}$ F/m, and $\mu = 2.5$ μH/m. Find: (a) E_y at $P(0, 0, 1.8$ cm$)$ at 0.2 ns; (b) H_x at P at 0.2 ns.

11.12 The plane wave $\mathbf{E_s} = 300e^{-jkx}\mathbf{a_y}$ V/m is propagating in a material for which $\mu = 2.25$ μH/m, $\epsilon' = 9$ pF/m, and $\epsilon'' = 7.8$ pF/m. If $\omega = 64$ Mrad/s, find: (a) α; (b) β; (c) v_p; (d) λ; (e) η; (f) $\mathbf{H_s}$; (g) $\mathbf{E}(3, 2, 4, 10$ ns$)$.

11.13 Let $jk = 0.2 + j1.5\ \text{m}^{-1}$ and $\eta = 450 + j60\Omega$ for a uniform plane propagating in the $\mathbf{a}_z$ direction. If $\omega = 300$ Mrad/s, find μ, ϵ', and ϵ'' for the medium.

11.14 A certain nonmagnetic material has the material constants $\epsilon'_R = 2$ and $\epsilon''/\epsilon' = 4 \times 10^{-4}$ at $\omega = 1.5$ Grad/s. Find the distance a uniform plane wave can propagate through the material before: (a) it is attenuated by 1 Np; (b) the power level is reduced by one-half; (c) the phase shifts 360°.

11.15 A 10-GHz radar signal may be represented as a uniform plane wave in a sufficiently small region. Calculate the wavelength in centimeters and the attenuation in nepers per meter if the wave is propagating in a nonmagnetic material for which: (a) $\epsilon'_R = 1$ and $\epsilon''_R = 0$; (b) $\epsilon'_R = 1.04$ and $\epsilon''_R = 9.00 \times 10^{-4}$; (c) $\epsilon'_R = 2.5$ and $\epsilon''_R = 7.2$.

11.16 The power factor of a capacitor is defined as the cosine of the impedance phase angle, and its Q is ωCR, where R is the parallel resistance. Assume an idealized parallel-plate capacitor having a dielectric characterized by σ, ϵ', and μ_R. Find both the power factor and Q in terms of the loss tangent.

11.17 Let $\eta = 250 + j30\Omega$ and $jk = 0.2 + j2\text{m}^{-1}$ for a uniform plane wave propagating in the $\mathbf{a}_z$ direction in a dielectric having some finite conductivity. If $|E_s| = 400$ V/m at $z = 0$, find: (a) $\mathcal{P}_{z,av}$ at $z = 0$ and $z = 60$ cm; (b) the average ohmic power dissipation in watts per cubic meter at $z = 60$ cm.

11.18 (a) Find $\mathcal{P}(\mathbf{r}, t)$ if $E_s = 400e^{-j2x}\mathbf{a}_y$ V/m in free space. (b) Find $\mathcal{P}$ at $t = 0$ for $\mathbf{r} = (a, 5, 10)$, where $a = 0, 1, 2$, and 3. (c) Find $\mathcal{P}$ at the origin for $t = 0, 0.2T, 0.4T$, and $0.6T$, where T is the oscillation period.

11.19 Perfectly conducting cylinders with radii of 8 mm and 20 mm are coaxial. The region between the cylinders is filled with a perfect dielectric for which $\epsilon = 10^{-9}/4\pi$ F/m and $\mu_R = 1$. If $\mathbf{E}$ in this region is $(500/\rho)\cos(\omega t - 4z)\mathbf{a}_\rho$ V/m, find: (a) ω, with the help of Maxwell's equations in cylindrical coordinates; (b) $\mathbf{H}(\rho, z, t)$; (c) $\mathcal{P}(\rho, z, t)$; (d) the average power passing through every cross section $8 < \rho < 20$ mm, $0 < \phi < 2\pi$.

11.20 If $\mathbf{E_s} = 60\frac{\sin\theta}{r}e^{-j2r}\mathbf{a}_\theta$ V/m and $\mathbf{H_s} = \frac{\sin\theta}{4\pi r}e^{-j2r}\mathbf{a}_\phi$ A/m in free space, find the average power passing outward through the surface $r = 10^6$, $0 < \theta < \pi/3$, $0 < \phi < 2\pi$.

11.21 The cylindrical shell, 1 cm $< \rho <$ 1.2 cm, is composed of a conducting material for which $\sigma = 10^6$ S/m. The external and internal regions are nonconducting. Let $H_\phi = 2000$ A/m at $\rho = 1.2$ cm. (a) Find $\mathbf{H}$ everywhere. (b) Find $\mathbf{E}$ everywhere. (c) Find $\mathcal{P}$ everywhere.

11.22 The inner and outer dimensions of a coaxial copper transmission line are 2 and 7 mm, respectively. Both conductors have thicknesses much greater than δ. The dielectric is lossless and the operating frequency is 400 MHz. Calculate the resistance per meter length of the: (a) inner conductor; (b) outer conductor; (c) transmission line.

11.23 A hollow tubular conductor is constructed from a type of brass having a conductivity of 1.2×10^7 S/m. The inner and outer radii are 9 and 10 mm, respectively. Calculate the resistance per meter length at a frequency of: (a) dc; (b) 20 MHz; (c) 2 GHz.

11.24 (a) Most microwave ovens operate at 2.45 GHz. Assume that $\sigma = 1.2 \times 10^6$ S/m and $\mu_R = 500$ for the stainless steel interior, and find the depth of penetration. (b) Let $E_s = 50\angle 0^\circ$ V/m at the surface of the conductor, and plot a curve of the amplitude of E_s vs. the angle of E_s as the field propagates into the stainless steel.

11.25 A good conductor is planar in form, and it carries a uniform plane wave that has a wavelength of 0.3 mm and a velocity of 3×10^5 m/s. Assuming the conductor is nonmagnetic, determine the frequency and the conductivity.

11.26 The dimensions of a certain coaxial transmission line are $a = 0.8$ mm and $b = 4$ mm. The outer conductor thickness is 0.6 mm, and all condcutors have $\sigma = 1.6 \times 10^7$ S/m. (a) Find R, the resistance per unit length at an operating frequency of 2.4 GHz. (b) Use information from Secs. 5.10 and 9.10 to find C and L, the capacitance and inductance per unit length, respectively. The coax is air-filled. (c) Find α and β if $\alpha + j\beta = \sqrt{j\omega C(R + j\omega L)}$.

11.27 The planar surface $z = 0$ is a brass-Teflon interface. Use data available in Appendix C to evaluate the following ratios for a uniform plane wave having $\omega = 4 \times 10^{10}$ rad/s: (a) $\alpha_{Tef}/\alpha_{brass}$; (b) $\lambda_{Tef}/\lambda_{brass}$; (c) v_{Tef}/v_{brass}.

11.28 A uniform plane wave in free space has electric field vector given by $\mathbf{E}_s = 10e^{-j\beta x}\mathbf{a}_z + 15e^{-j\beta x}\mathbf{a}_y$ V/m. (a) Describe the wave polarization: (b) Find $\mathbf{H}_s$; (c) determine the average power density in the wave in W/m^2.

11.29 Consider a left-circularly polarized wave in free space that propagates in the forward z direction. The electric field is given by the appropriate form of Eq. (80). (a) Determine the magnetic field phasor, $\mathbf{H}_s$; (b) determine an expression for the average power density in the wave in W/m^2 by direct application of Eq. (57).

11.30 The electric field of a uniform plane wave in free space is given by $\mathbf{E}_s = 10(\mathbf{a}_y + j\mathbf{a}_z)e^{-j25x}$. (a) Determine the frequency, f; (b) find the magnetic field phasor, $\mathbf{H}_s$; (c) describe the polarization of the wave.

11.31 A linearly polarized uniform plane wave, propagating in the forward z direction, is input to a lossless *anisotropic* material, in which the dielectric constant encountered by waves polarized along $y(\epsilon_{Ry})$ differs from that seen by waves polarized along $x(\epsilon_{Rx})$. Suppose $\epsilon_{Rx} = 2.15$, $\epsilon_{Ry} = 2.10$, and the wave electric field at input is polarized at 45° to the positive x and y axes. (a) Determine the shortest length of the material such that the wave as it emerges from the output end is circularly polarized; (b) will the output wave be right- or left-circularly polarized?

11.32 Suppose that the length of the medium of Problem 11.31. is made to be *twice* that as determined in the problem. Describe the polarization of the output wave in this case.

11.33 Given a wave for which $\mathbf{E}_s = 15e^{-j\beta z}\mathbf{a}_x + 18e^{-j\beta z}e^{j\phi}\mathbf{a}_y$ V/m: (a) Find $\mathbf{H}_s$; (b) determine the average power density in W/m^2

11.34 Given a general elliptically polarized wave as per Eq. (73):

$$\mathbf{E}_s = [E_{x0}\mathbf{a}_x + E_{y0}e^{j\phi}\mathbf{a}_y]e^{-j\beta z}$$

(a) Show, using methods similar to those of Example 11.7, that a linearly polarized wave results when superimposing the given field and a phase-shifted field of the form:

$$\mathbf{E}_s = [E_{x0}\mathbf{a}_x + E_{y0}e^{-j\phi}\mathbf{a}_y]e^{-j\beta z}e^{j\delta}$$

where δ is a constant; (b) Find δ in terms of ϕ such that the resultant wave is linearly polarized along x.

CHAPTER

12

PLANE WAVES AT BOUNDARIES AND IN DISPERSIVE MEDIA

In Chapter 11, we considered basic electromagnetic wave principles. We learned how to mathematically represent waves as functions of frequency, medium properties, and electric field orientation. We also learned how to calculate the wave velocity, attenuation, and power. In this chapter, we consider wave reflection and transmission at planar boundaries between media having different properties. Our study will allow any orientation between the wave and boundary, and will also include the important cases of multiple boundaries. We will also study the practical case of waves that carry power over a finite band of frequencies, as would occur, for example, in a modulated carrier. We will consider such waves in *dispersive* media, in which some parameter that affects propagation (permittivity for example) varies with frequency. The effect of a dispersive medium on a signal is of great importance, since the signal envelope will change its shape as it propagates. This can occur to an extent that at the receiving end, detection and faithful representation of the original signal become problematic. Dispersion thus becomes an important limiting factor in allowable propagation distances, in a similar way that we found to be true for attenuation.

12.1 REFLECTION OF UNIFORM PLANE WAVES AT NORMAL INCIDENCE

In this section we will consider the phenomenon of reflection which occurs when a uniform plane wave is incident on the boundary between regions composed of two different materials. The treatment is specialized to the case of *normal incidence*, in which the wave propagation direction is perpendicular to the boundary. In later sections we remove this restriction. We shall establish expressions for the wave that is reflected from the interface and for that which is transmitted from one region into the other. These results will be directly applicable to impedance-matching problems in ordinary transmission lines as well as to waveguides and other more exotic transmission systems.

We again assume that we have only a single vector component of the electric field intensity. Referring to Fig. 12.1, we define region 1 (ϵ_1, μ_1) as the half-space for which $z < 0$; region 2 (ϵ_2, μ_2) is the half-space for which $z > 0$. Initially we establish the wave traveling in the $+z$ direction in region 1,

$$E_{x1}^{+}(z, t) = E_{x10}^{+}\, e^{-\alpha_1 z} \cos(\omega t - \beta_1 z)$$

or

$$E_{xs1}^{+} = E_{x10}^{+}\, e^{-jk_1 z} \tag{1}$$

where we take E_{x10}^{+} as real. The subscript 1 identifies the region and the superscript $+$ indicates a positively traveling wave. Associated with $E_{x1}^{+}(z, t)$ is a magnetic field

$$H_{ys1}^{+} = \frac{1}{\eta_1} E_{x10}^{+}\, e^{-jk_1 z} \tag{2}$$

where k_1 and η_1 are complex unless ϵ_1'' (or σ_1) is zero. This uniform plane wave in region 1 which is traveling toward the boundary surface at $z = 0$ is called the *incident* wave. Since the direction of propagation of the incident wave is perpendicular to the boundary plane, we describe it as normal incidence.

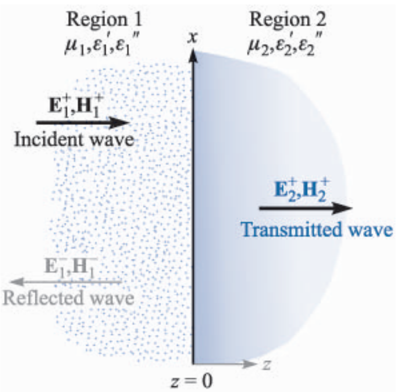


FIGURE 12.1
A wave $\mathbf{E}_1^+$ incident on a plane boundary establishes a reflected wave $\mathbf{E}_1^-$ and a transmitted wave $\mathbf{E}_2^+$.

We now recognize that energy may be transmitted across the boundary surface at $z = 0$ into region 2 by providing a wave moving in the $+z$ direction in that medium,

$$E^+_{xs2} = E^+_{x20}\, e^{-jk_2 z} \tag{3}$$

$$H^+_{ys2} = \frac{1}{\eta_2} E^+_{x20}\, e^{-jk_2 z} \tag{4}$$

This wave which moves away from the boundary surface into region 2 is called the *transmitted* wave; note the use of the different propagation constant k_2 and intrinsic impedance η_2.

Now we must try to satisfy the boundary conditions at $z = 0$ with these assumed fields. E_x is a tangential field; therefore the **E** fields in regions 1 and 2 must be equal at $z = 0$. Setting $z = 0$ in (1) and (3) would require that $E^+_{x10} = E^+_{x20}$. H_y is also a tangential field, however, and must be continuous across the boundary (no current sheets are present in real media). When we let $z = 0$ in (2) and (4), however, we find that we must have $E^+_{x10}/\eta_1 = E^+_{x20}/\eta_2$; since $E^+_{x10} = E^+_{x20}$, then $\eta_1 = \eta_2$. But this is a very special condition that does not fit the facts in general, and we are therefore unable to satisfy the boundary conditions with only an incident and a transmitted wave. We require a wave traveling away from the boundary in region 1, as shown in Fig. 12.1; this is called a *reflected* wave,

$$E^-_{xs1} = E^-_{x10}\, e^{jk_1 z} \tag{5}$$

$$H^-_{ys1} = -\frac{E^-_{x10}}{\eta_1}\, e^{jk_1 z} \tag{6}$$

where E^-_{x10} may be a complex quantity. Since this field is traveling in the $-z$ direction, $E^-_{xs1} = -\eta_1 H^-_{ys1}$, for the Poynting vector shows that $\mathbf{E}^-_1 \times \mathbf{H}^-_1$ must be in the $-\mathbf{a_z}$ direction.

The boundary conditions are now easily satisfied, and in the process the amplitudes of the transmitted and reflected waves may be found in terms of E^+_{x10}. The total electric field intensity is continuous at $z = 0$,

$$E_{xs1} = E_{xs2} \qquad (z = 0)$$

or

$$E^+_{xs1} + E^-_{xs1} = E^+_{xs2} \qquad (z = 0)$$

Therefore

$$\boxed{E^+_{x10} + E^-_{x10} = E^+_{x20}} \tag{7}$$

Furthermore,

$$H_{ys1} = H_{ys2} \qquad (z = 0)$$

or

$$H_{ys1}^{+} + H_{ys1}^{-} = H_{ys2}^{+} \qquad (z = 0)$$

and therefore

$$\boxed{\frac{E_{x10}^{+}}{\eta_1} - \frac{E_{x10}^{-}}{\eta_1} = \frac{E_{x20}^{+}}{\eta_2}} \tag{8}$$

Solving (8) for E_{x20}^{+} and substituting into (7), we find

$$E_{x10}^{+} + E_{x10}^{-} = \frac{\eta_2}{\eta_1} E_{x10}^{+} - \frac{\eta_2}{\eta_1} E_{x10}^{-}$$

or

$$E_{x10}^{-} = E_{x10}^{+} \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1}$$

The ratio of the amplitudes of the reflected and incident electric fields is called the *reflection coefficient* and is designated by Γ (gamma),

$$\boxed{\Gamma = \frac{E_{x10}^{-}}{E_{x10}^{+}} = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1}} \tag{9}$$

The reflection coefficient may be complex, in which case there is a phase shift in the reflected wave.

The relative amplitude of the transmitted electric field intensity is found by combining (9) and (7), to yield the *transmission coefficient*, τ:

$$\boxed{\tau = \frac{E_{x20}^{+}}{E_{x10}^{+}} = \frac{2\eta_2}{\eta_1 + \eta_2} = 1 + \Gamma} \tag{10}$$

Let us see how these results may be applied to several special cases. We first let region 1 be a perfect dielectric and region 2 be a perfect conductor. Then, since σ_2 is infinite,

$$\eta_2 = \sqrt{\frac{j\omega\mu_2}{\sigma_2 + j\omega\epsilon_2'}} = 0$$

and from (10),

$$E_{x20}^{+} = 0$$

No time-varying fields can exist in the *perfect* conductor. An alternate way of looking at this is to note that the skin depth is zero.

Since $\eta_2 = 0$, then (9) shows that

$$\Gamma = -1$$

and

$$E_{x10}^{+} = -E_{x10}^{-}$$

The incident and reflected fields are of equal amplitude, and so all the incident energy is reflected by the perfect conductor. The fact that the two fields are of opposite sign indicates that at the boundary (or at the moment of reflection) the reflected field is shifted in phase by 180° relative to the incident field. The total **E** field in region 1 is

$$\begin{aligned} E_{xs1} &= E_{xs1}^{+} + E_{xs1}^{-} \\ &= E_{x10}^{+} e^{-j\beta_1 z} - E_{x10}^{+} e^{j\beta_1 z} \end{aligned}$$

where we have let $jk_1 = 0 + j\beta_1$ in the perfect dielectric. These terms may be combined and simplified,

$$\begin{aligned} E_{xs1} &= \left(e^{-j\beta_1 z} - e^{j\beta_1 z}\right) E_{x10}^{+} \\ &= -j2 \sin(\beta_1 z)\, E_{x10}^{+} \end{aligned} \tag{11}$$

Multiplying (11) by $e^{j\omega t}$ and taking the real part, we may drop the s subscript and obtain the real instantaneous form:

$$\boxed{E_{x1}(z, t) = 2E_{x10}^{+} \sin(\beta_1 z) \sin(\omega t)} \tag{12}$$

This total field in region 1 is not a traveling wave, although it was obtained by combining two waves of equal amplitude traveling in opposite directions. Let us compare its form with that of the incident wave,

$$E_{x1}(z, t) = E_{x10}^{+} \cos(\omega t - \beta_1 z) \tag{13}$$

Here we see the term $\omega t - \beta_1 z$ or $\omega(t - z/v_{p1})$, which characterizes a wave traveling in the $+z$ direction at a velocity $v_{p1} = \omega/\beta_1$. In (12), however, the factors involving time and distance are separate trigonometric terms. At all planes for which $\beta_1 z = m\pi$, E_{x1} is zero for all time. Furthermore, whenever $\omega t = m\pi$, E_{x1} is zero everywhere. A field of the form of (12) is known as a *standing wave*.

The planes on which $E_{x1} = 0$ are located where

$$\beta_1 z = m\pi \qquad (m = 0, \pm 1, \pm 2, \ldots)$$

Thus

$$\frac{2\pi}{\lambda_1} z = m\pi$$

and

$$z = m\frac{\lambda_1}{2}$$

Thus $E_{x1} = 0$ at the boundary $z = 0$ and at every half-wavelength from the boundary in region 1, $z < 0$, as illustrated in Fig. 12.2.

Since $E^+_{xs1} = \eta_1 H^+_{ys1}$ and $E^-_{xs1} = -\eta_1 H^-_{ys1}$, the magnetic field is

$$H_{ys1} = \frac{E^+_{x10}}{\eta_1}\left(e^{-j\beta_1 z} + e^{j\beta_1 z}\right)$$

or

$$\boxed{H_{y1}(z, t) = 2\frac{E^+_{x10}}{\eta_1}\cos(\beta_1 z)\cos(\omega t)} \tag{14}$$

This is also a standing wave, but it shows a maximum amplitude at the positions where $E_{x1} = 0$. It is also 90° out of time phase with E_{x1} everywhere. Thus no average power is transmitted in either direction.

Let us now consider perfect dielectrics in both regions 1 and 2; η_1 and η_2 are both real positive quantities and $\alpha_1 = \alpha_2 = 0$. Equation (9) enables us to calcu-

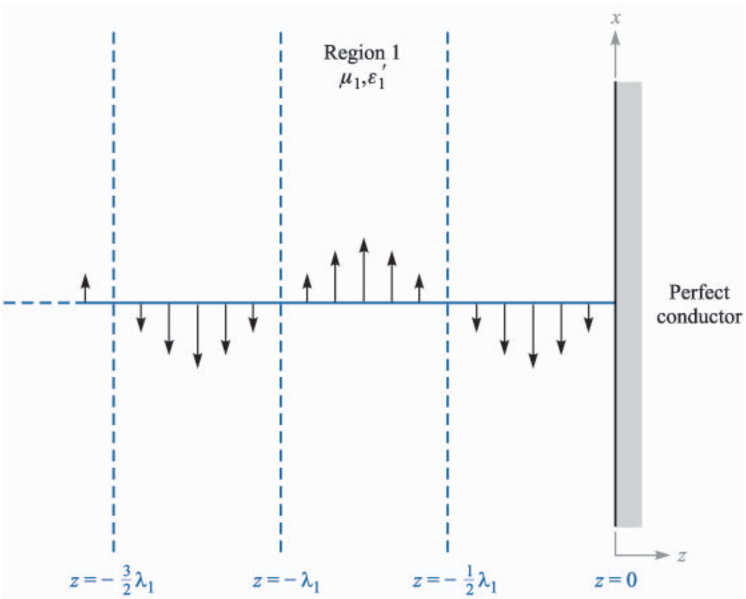


FIGURE 12.2
The instantaneous values of the total field E_{x1} are shown at $t = \pi/2$. $E_{x1} = 0$ for all time at multiples of one half-wavelength from the conducting surface.

late the reflection coefficient and find E_{x1}^- in terms of the incident field E_{x1}^+. Knowing E_{x1}^+ and E_{x1}^-, we then find H_{y1}^+ and H_{y1}^-. In region 2, E_{x2}^+ is found from (10), and this then determines H_{y2}^+.

Example 12.1

As a numerical example let us select

$$\eta_1 = 100 \quad \Omega$$
$$\eta_2 = 300 \quad \Omega$$
$$E_{x10}^+ = 100 \quad \text{V/m}$$

and calculate values for the incident, reflected, and transmitted waves.

Solution. The reflection coefficient is

$$\Gamma = \frac{300 - 100}{300 + 100} = 0.5$$

and thus

$$E_{x10}^- = 50 \quad \text{V/m}$$

The magnetic field intensities are

$$H_{y10}^+ = \frac{100}{100} = 1.00 \quad \text{A/m}$$

$$H_{y10}^- = -\frac{50}{100} = -0.50 \quad \text{A/m}$$

The incident power density is

$$\mathcal{P}_{1,av}^+ = \frac{1}{2} E_{x10}^+ H_{y10}^+ = 100 \quad \text{W/m}^2$$

while

$$\mathcal{P}_{1,av}^- = -\frac{1}{2} E_{x10}^- H_{y10}^- = 25.0 \quad \text{W/m}^2$$

In region 2, using (10)

$$E_{x20}^+ = \tau E_{x10}^+ = 150 \quad \text{V/m}$$

and

$$H_{y20}^+ = \frac{150}{300} = 0.500 \quad \text{A/m}$$

Thus

$$\mathcal{P}_{2,av}^+ = \frac{1}{2} E_{x20}^+ H_{y20}^+ = 75.0 \quad \text{W/m}^2$$

Note that energy is conserved:

$$\mathcal{P}_{1,av}^+ = \mathcal{P}_{1,av}^- + \mathcal{P}_{2,av}^+$$

We can formulate a general rule on the transfer of power through reflection and transmission by using Eq. (57) from Chapter 11:

$$\mathcal{P}_{av} = \frac{1}{2}\text{Re}\{\mathbf{E}_s \times \mathbf{H}_s^*\}$$

We consider the same field vector and interface orientations as before, but we consider the general case of complex impedances. For the incident power density, we have

$$\mathcal{P}_{1,av}^{+} = \frac{1}{2}\text{Re}\left\{E_{x10}^{+} H_{y10}^{+*}\right\} = \frac{1}{2}\text{Re}\left\{E_{x10}^{+} \frac{1}{\eta_1^*} E_{x10}^{+*}\right\} = \frac{1}{2}\text{Re}\left\{\frac{1}{\eta_1^*}\right\}|E_{x10}^{+}|^2$$

The reflected power density is then

$$\mathcal{P}_{1,av}^{-} = -\frac{1}{2}\text{Re}\left\{E_{x10}^{-} H_{y10}^{-*}\right\} = \frac{1}{2}\text{Re}\left\{\Gamma E_{x10}^{+} \frac{1}{\eta_1^*} \Gamma^* E_{x10}^{+*}\right\} = \frac{1}{2}\text{Re}\left\{\frac{1}{\eta_1^*}\right\}|E_{x10}^{+}|^2\,|\Gamma|^2$$

We thus find the general relation between the reflected and incident power:

$$\boxed{\mathcal{P}_{1,av}^{-} = |\Gamma|^2 \mathcal{P}_{1,av}^{+}} \tag{15}$$

In a similar way, we find the transmitted power:

$$\mathcal{P}_{2,av}^{+} = \frac{1}{2}\text{Re}\left\{E_{x20}^{+} H_{y20}^{+*}\right\} = \frac{1}{2}\text{Re}\left\{\tau E_{x10}^{+} \frac{1}{\eta_2^*} \tau^* E_{x10}^{+*}\right\} = \frac{1}{2}\text{Re}\left\{\frac{1}{\eta_2^*}\right\}|E_{x10}^{+}|^2\,|\tau|^2$$

and so we see that the incident and transmitted powers are related through

$$\mathcal{P}_{2,av}^{+} = \frac{\text{Re}\{1/\eta_2^*\}}{\text{Re}\{1/\eta_1^*\}}\,|\tau|^2\,\mathcal{P}_{1,av}^{+} = \left|\frac{\eta_1}{\eta_2}\right|^2 \left(\frac{\eta_2 + \eta_2^*}{\eta_1 + \eta_1^*}\right)|\tau|^2\,\mathcal{P}_{1,av}^{+} \tag{16}$$

Eq. (16) is a relatively complicated way to calculate the transmitted power, unless the impedances are real. It is easier to take advantage of energy conservation by noting that whatever power is not reflected must be transmitted. Eq. (15) can thus be used to find

$$\boxed{\mathcal{P}_{2,av}^{+} = \left(1 - |\Gamma|^2\right)\mathcal{P}_{1,av}^{+}} \tag{17}$$

As would be expected (and which must be true), Eq. (17) can also be derived from Eq. (16).

✔ **D12.1.** A 1 MHz uniform plane wave is normally incident onto a freshwater lake ($\epsilon_R' = 78$, $\epsilon_R'' = 0$, $\mu_R = 1$). Determine the fraction of the incident power that is (a) reflected and (b) transmitted; (c) determine the amplitude of the electric field that is transmitted into the lake.

Ans. 0.63; 0.37; 0.20 V/m.

12.2 STANDING WAVE RATIO

One of the measurements that is easily made on transmission systems is the *relative* amplitude of the electric or magnetic field intensity through use of a probe. A small coupling loop will give an indication of the amplitude of the magnetic field, while a slightly extended center conductor of a coaxial cable will sample the electric field. Both devices are customarily tuned to the operating frequency to provide increased sensitivity. The output of the probe is rectified and connected directly to a microammeter, or it may be delivered to an electronic voltmeter or a special amplifier. The indication is proportional to the amplitude of the sinusoidal time-varying field in which the probe is immersed.

When a uniform plane wave is traveling through a lossless region, and no reflected wave is present, the probe will indicate the same amplitude at every point. Of course, the instantaneous field which the probe samples will differ in phase by $\beta(z_2 - z_1)$ rad as the probe is moved from $z = z_1$ to $z = z_2$, but the system is insensitive to the phase of the field. The equal-amplitude voltages are characteristic of an unattenuated traveling wave.

When a wave traveling in a lossless medium is reflected by a perfect conductor, the total field is a standing wave and, as shown by Eq. (12), the voltage probe provides no output when it is located an integral number of half-wavelengths from the reflecting surface. As the probe position is changed, its output varies as $|\sin \beta z|$, where z is the distance from the conductor. This sinusoidal amplitude variation is shown in Fig. 12.3, and it characterizes a standing wave.

A more complicated situation arises when the reflected field is neither 0 nor 100 percent of the incident field. Some energy is transmitted into the second region and some is reflected. Region 1 therefore supports a field that is composed of both a traveling wave and a standing wave. It is customary to describe this field as a standing wave even though a traveling wave is also present. We shall see that the field does not have zero amplitude at any point for all time, and the degree to which the field is divided between a traveling wave and a true standing wave is expressed by the ratio of the maximum amplitude found by the probe to the minimum amplitude.

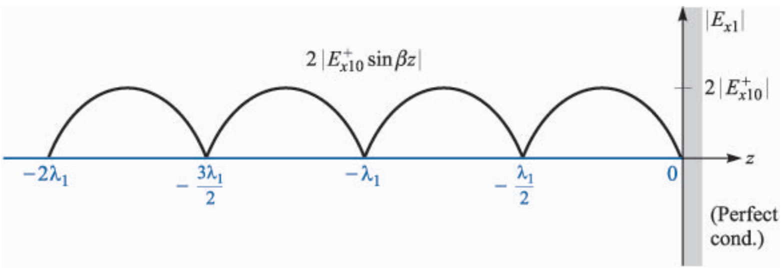


FIGURE 12.3
The standing voltage wave produced in a lossless medium by reflection from a perfect conductor varies as $|\sin \beta z|$.

Using the same fields investigated in the previous section, we combine the incident and reflected electric field intensities,

$$E_{x1} = E_{x1}^{+} + E_{x1}^{-}$$

The field E_{x1} is a sinusoidal function of t (generally with a nonzero phase angle), and it varies with z in a manner as yet unknown. We shall inspect all z to find the maximum and minimum amplitudes, and determine their ratio. We call this ratio the *standing-wave ratio*, and we shall symbolize it by s.

Let us now go through the mechanics of this procedure for the case in which medium 1 is a perfect dielectric, $\alpha_1 = 0$, but region 2 may be any material. We have

$$E_{xs1}^{+} = E_{x10}^{+}\, e^{-j\beta_1 z}$$

$$E_{xs1}^{-} = \Gamma E_{x10}^{+}\, e^{j\beta_1 z}$$

where

$$\Gamma = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1}$$

and η_1 is real and positive but η_2 may be complex. Thus Γ may be complex, and we allow for this possibility by letting

$$\Gamma = |\Gamma|\, e^{j\phi}$$

If region 2 is a perfect conductor, ϕ is equal to π; if η_2 is real and less than η_1, ϕ is also equal to π; and if η_2 is real and greater than η_1, ϕ is zero. The total field in region 1 is

$$E_{xs1} = \left(e^{-j\beta_1 z} + |\Gamma| e^{j(\beta_1 z + \phi)}\right) E_{x10}^{+} \tag{18}$$

We seek the maximum and minimum values of the magnitude of the complex quantity in the larger parentheses in (18). We certainly have a maximum when each term in the larger parentheses has the same phase angle; thus, for E_{x10}^{+} positive and real,

$$E_{xs1,\,max} = (1 + |\Gamma|) E_{x10}^{+} \tag{19}$$

and this occurs where

$$-\beta_1 z = \beta_1 z + \phi + 2m\pi \qquad (m = 0,\ \pm 1,\ \pm 2, \ldots) \tag{20}$$

Thus

$$\boxed{z_{max} = -\frac{1}{2\beta_1}(\phi + 2m\pi)} \tag{21}$$

Note that a field maximum is located at the boundary plane ($z = 0$) if $\phi = 0$; moreover, $\phi = 0$ when Γ is real and positive. This occurs for real η_1 and η_2 when

$\eta_2 > \eta_1$. Thus there is a voltage maximum at the boundary surface when the intrinsic impedance of region 2 is greater than that of region 1 and both impedances are real. With $\phi = 0$, maxima also occur at $z_{max} = -m\pi/\beta_1 = -m\lambda_1/2$.

For the perfect conductor $\phi = \pi$, and these maxima are found at $z_{max} = -\pi/(2\beta_1)$, $-3\pi/(2\beta_1)$, or $z_{max} = -\lambda_1/4$, $-3\lambda_1/4$, and so forth.

The minima must occur where the phase angles of the two terms in the larger parentheses in (18) differ by 180°; thus

$$E_{xs1,\,min} = (1 - |\Gamma|)E^+_{x10} \tag{22}$$

and this occurs where

$$-\beta_1 z = \beta_1 z + \phi + \pi + 2m\pi \qquad (m = 0, \pm 1, \pm 2, \ldots) \tag{23}$$

or

$$\boxed{z_{min} = -\frac{1}{2\beta_1}(\phi + (2m+1)\pi)} \tag{24}$$

The minima are separated by multiples of one half-wavelength (as are the maxima), and for the perfect conductor the first minimum occurs when $-\beta_1 z = 0$, or at the conducting surface. In general, an electric field minimum is found at $z = 0$ whenever $\phi = \pi$; this occurs if $\eta_2 < \eta_1$ and both are real. The general results are illustrated in Fig. 12.4.

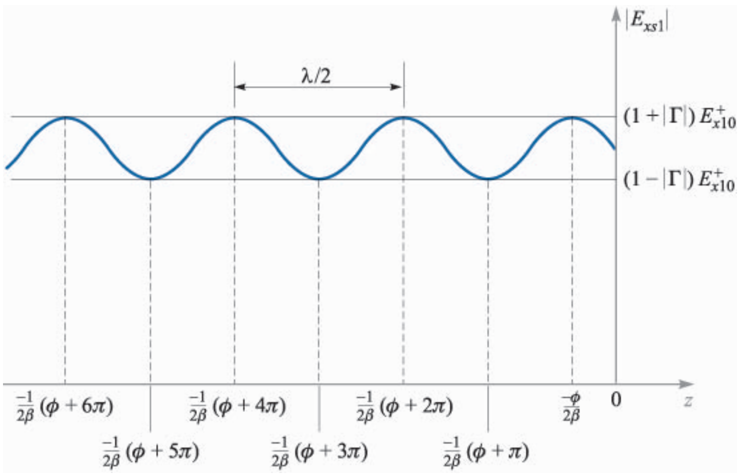


FIGURE 12.4
Plot of the magnitude of E_{xs1} as found from Eq. (18) as a function of position, z, in front of the interface (at $z = 0$). The reflection coefficient phase is ϕ, which leads to the indicated locations of maximum and minimum **E**, as found through Eqs. (21) and (24).

Further insights can be obtained by revisiting Eq. (18), which can be rewritten in the form

$$E_{xs1} = E_{x10}^{+}\left(e^{-j\phi/2}e^{-j\beta_1 z} + |\Gamma|e^{j\phi/2}e^{j\beta_1 z}\right)e^{j\phi/2}$$

As a trick, we can add and subtract $\left(|\Gamma|E_{x10}^{+}\, e^{-j\phi/2}e^{-j\beta_1 z}\right)$ to obtain

$$E_{xs1} = E_{x10}^{+}(1-|\Gamma|)e^{-j\beta_1 z} + E_{x10}^{+}|\Gamma|\left(e^{-j\phi/2}e^{-j\beta_1 z} + e^{j\phi/2}e^{j\beta_1 z}\right)e^{j\phi/2}$$

which reduces to

$$E_{xs1} = (1-|\Gamma|)E_{x10}^{+}e^{-j\beta_1 z} + 2|\Gamma|E_{x10}^{+}\, e^{j\phi/2}\cos(\beta_1 z + \phi/2) \tag{25}$$

Finally, the real instantaneous form of (25) is obtained through $E_{x1}(z, t) = \text{Re}\left\{E_{xs1}e^{j\omega t}\right\}$. We find

$$\boxed{\begin{aligned} E_{x1}(z,t) = {} & (1-|\Gamma|)E_{x10}^{+}\cos(\omega t - \beta_1 z) \\ & + 2|\Gamma|E_{x10}^{+}\cos(\beta_1 z + \phi/2)\cos(\omega t + \phi/2) \end{aligned}} \tag{26}$$

Equation (26) is recognized as the sum of a traveling wave of amplitude $(1-|\Gamma|)E_{x10}^{+}$ and a standing wave having amplitude $2|\Gamma|E_{x10}^{+}$. We can visualize events as follows: The portion of the incident wave that reflects and back-propagates in region 1 interferes with an equivalent portion of the incident wave to form a standing wave. The rest of the incident wave (that does not interfere) is the traveling wave part of (26). The maximum amplitude observed in region 1 is found where the amplitudes of the two terms in (26) add directly to give $(1+|\Gamma|)E_{x10}^{+}$. The minimum amplitude is found where the standing wave achieves a null, leaving only the traveling wave amplitude of $(1-|\Gamma|)E_{x10}^{+}$. The fact that the two terms in (26) combine in this way with the proper phasing is not readily apparent, but can be confirmed by substituting z_{max} and z_{min}, as given by (21) and (24).

Example 12.2

To illustrate some of these results, let us consider a 100-V/m, 3-GHz wave that is propagating in a material having $\epsilon'_{R1} = 4$, $\mu_{R1} = 1$, and $\epsilon''_R = 0$. The wave is normally incident on another perfect dielectric in region 2, $z > 0$, where $\epsilon'_{R2} = 9$ and $\mu_{R2} = 1$ (Fig. 12.5). We seek the locations of the maxima and minima of **E**.

Solution. We calculate $\omega = 6\pi \times 10^9$ rad/s, $\beta_1 = \omega\sqrt{\mu_1\epsilon_1} = 40\pi$ rad/m, and $\beta_2 = \omega\sqrt{\mu_2\epsilon_2} = 60\pi$ rad/m. Although the wavelength would be 10 cm in air, we find here that $\lambda_1 = 2\pi/\beta_1 = 5$ cm, $\lambda_2 = 2\pi/\beta_2 = 3.33$ cm, $\eta_1 = 60\pi\,\Omega$, $\eta_2 = 40\pi\,\Omega$, and $\Gamma = (\eta_2 - \eta_1)/(\eta_2 + \eta_1) = -0.2$. Since Γ is real and negative ($\eta_2 < \eta_1$), there will be a minimum of the electric field at the boundary, and it will be repeated at half-wavelength (2.5 cm) intervals in dielectric 1. From (22), we see that $E_{xs1,\,min} = 80$ V/m.

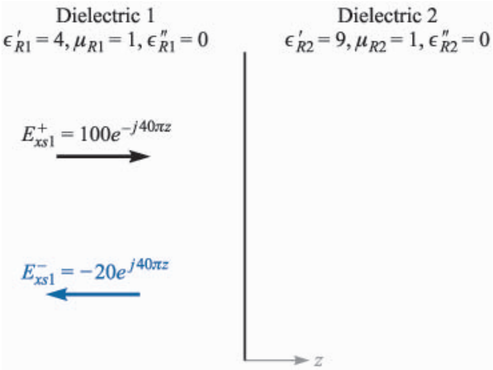


FIGURE 12.5
An incident wave, $E^+_{xs1} = 100e^{-j40\pi z}$ V/m is reflected with a reflection coefficient $\Gamma = -0.2$. Dielectric 2 is infinitely thick.

Maxima of **E** are found at distances of 1.25, 3.75, 6.25, ... cm from $z = 0$. These maxima all have amplitudes of 120 V/m, as predicted by (19).

There are no maxima or minima in region 2 since there is no reflected wave there.

The ratio of the maximum to minimum amplitudes is called the *standing wave ratio*:

$$s = \frac{E_{xs1,\,max}}{E_{xs1,\,min}} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{27}$$

Since $|\Gamma| < 1$, s is always positive and greater than or equal to unity. For the example above,

$$s = \frac{1 + |-0.2|}{1 - |-0.2|} = \frac{1.2}{0.8} = 1.5$$

If $|\Gamma| = 1$, the reflected and incident amplitudes are equal, all the incident energy is reflected, and s is infinite. Planes separated by multiples of $\lambda_1/2$ can be found on which E_{x1} is zero at all times. Midway between these planes, E_{x1} has a maximum amplitude twice that of the incident wave.

If $\eta_2 = \eta_1$, then $\Gamma = 0$, no energy is reflected, and $s = 1$; the maximum and minimum amplitudes are equal.

If one-half the incident power is reflected, $|\Gamma|^2 = 0.5$, $|\Gamma| = 0.707$, and $s = 5.83$.

✔ **D12.2.** What value of s results when $\Gamma = \pm 1/2$?

Ans. 3

Since the standing-wave ratio is a ratio of amplitudes, the relative amplitudes provided by a probe permit its use to determine s experimentally.

Example 12.3

A uniform plane wave in air partially reflects from the surface of a material whose properties are unknown. Measurements of the electric field in the region in front of the interface yield a 1.5 m spacing between maxima, with the first maximum occurring 0.75 m from the interface. A standing wave ratio of 5 is measured. Determine the intrinsic impedance, η_u, of the unknown material.

Solution. The 1.5 m spacing between maxima is $\lambda/2$, implying a wavelength is 3.0 m, or $f = 100$ MHz. The first maximum at 0.75 m is thus at a distance of $\lambda/4$ from the interface, which means that a field minimum occurs at the boundary. Thus Γ will be real and negative. We use (27) to write

$$|\Gamma| = \frac{s-1}{s+1} = \frac{5-1}{5+1} = \frac{2}{3}$$

So

$$\Gamma = -\frac{2}{3} = \frac{\eta_u - \eta_0}{\eta_u + \eta_0}$$

which we solve for η_u to obtain

$$\eta_u = \frac{1}{5}\eta_0 = \frac{377}{5} = 75.4\,\Omega$$

12.3 WAVE REFLECTION FROM MULTIPLE INTERFACES

So far we have treated the reflection of waves at the single boundary that occurs between semi-infinite media. In this section, we consider wave reflection from materials that are finite in extent, such that we must consider the effect of the front and back surfaces. Such a two-interface problem would occur, for example, for light incident on a flat piece of glass. Additional interfaces are present if the glass is coated with one or more layers of dielectric material for the purpose (as we will see) of reducing reflections. Such problems in which more than one interface is involved are frequently encountered; single interface problems are in fact more the exception than the rule.

Consider the general situation shown in Fig. 12.6, in which a uniform plane wave propagating in the forward z direction is normally incident from the left onto the interface between regions 1 and 2; these have intrinsic impedances, η_1 and η_2. A third region of impedance η_3 lies beyond region 2, and so a second interface exists between regions 2 and 3. We let the second interface location occur at $z = 0$, and so all positions to the left will be described by values of z that are negative. The width of the second region is l, so the first interface will occur at position $z = -l$.

When the incident wave reaches the first interface, events occur as follows: A portion of the wave reflects, while the remainder is transmitted, to propagate toward the second interface. There, a portion is transmitted into region 3, while

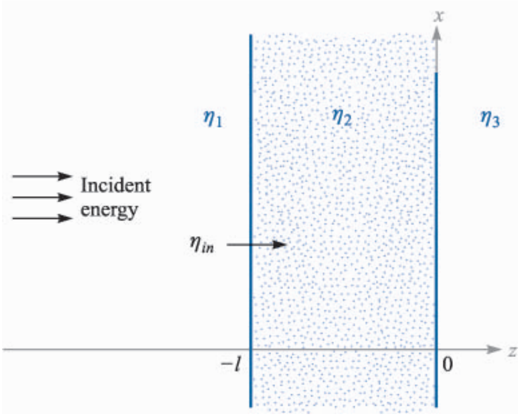


FIGURE 12.6
Basic two-interface problem, in which the impedances of regions 2 and 3, along with the finite thickness of region 2, are accounted for in the input impedance at the front surface, η_{in}.

the rest reflects and returns to the first interface; there it is again partially reflected. This reflected wave then combines with additional transmitted energy from region 1, and the process repeats. We thus have a complicated sequence of multiple reflections that occur within region 2, with partial transmission at each bounce. To analyze the situation in this way would involve keeping track of a very large number of reflections; this would be necessary when studying the *transient* phase of the process, where the incident wave first encounters the interfaces.

If the incident wave is left on for all time, however, a *steady-state* situation is eventually reached, in which: (1) an overall fraction of the incident wave is reflected from the two-interface configuration and back-propagates in region 1 with a definite amplitude and phase; (2) an overall fraction of the incident wave is transmitted through the two interfaces and forward-propagates in the third region; (3) a net backward wave exists in region 2, consisting of all reflected waves from the second interface; (4) a net forward wave exists in region 2, which is the superposition of the transmitted wave through the first interface, and all waves in region 2 that have reflected from the first interface and are now forward-propagating. The effect of combining many co-propagating waves in this way is to establish a single wave which has a definite amplitude and phase, determined through the sums of the amplitudes and phases of all the component waves. In steady state, we thus have a total of five waves to consider. These are the incident and net reflected waves in region 1, the net transmitted wave in region 3, and the two counter-propagating waves in region 2.

Let us assume all regions are composed of lossless media, and consider the two waves in region 2. Taking these as x-polarized, their electric fields add to yield

$$E_{xs2} = E^{+}_{x20}\, e^{-j\beta_2 z} + E^{-}_{x20}\, e^{j\beta_2 z} \tag{28}$$

where $\beta_2 = \omega\sqrt{\epsilon_{R2}}/c$, and where the amplitudes, E^{+}_{x20} and E^{-}_{x20}, are complex. The y-polarized magnetic field is similarly written, using complex amplitudes:

$$H_{ys2} = H^{+}_{y20}\, e^{-j\beta_2 z} + H^{-}_{y20}\, e^{j\beta_2 z} \tag{29}$$

We now note that the forward and backward electric field amplitudes in region 2 are related through the reflection coefficient at the second interface, Γ_{23}, where

$$\Gamma_{23} = \frac{\eta_3 - \eta_2}{\eta_3 + \eta_2} \tag{30}$$

We thus have

$$E_{x20}^- = \Gamma_{23} E_{x20}^+ \tag{31}$$

We then write the magnetic field amplitudes in terms of electric field amplitudes through

$$H_{y20}^+ = \frac{1}{\eta_2} E_{x20}^+ \tag{32}$$

and

$$H_{y20}^- = -\frac{1}{\eta_2} E_{x20}^- = -\frac{1}{\eta_2} \Gamma_{23} E_{x20}^+ \tag{33}$$

We now define the *wave impedance*, η_w, as the z-dependent ratio of the total electric field to the total magnetic field. In region 2, this becomes, using (28) and (29)

$$\eta_w(z) = \frac{E_{xs2}}{H_{ys2}} = \frac{E_{x20}^+ e^{-j\beta_2 z} + E_{x20}^- e^{j\beta_2 z}}{H_{y20}^+ e^{-j\beta_2 z} + H_{y20}^- e^{j\beta_2 z}}$$

Then, using (31), (32), and (33), we obtain

$$\eta_w(z) = \eta_2 \left[\frac{e^{-j\beta_2 z} + \Gamma_{23} e^{j\beta_2 z}}{e^{-j\beta_2 z} - \Gamma_{23} e^{j\beta_2 z}} \right] \tag{34}$$

Now, using (30) and Euler's identity, we have

$$\eta_w(z) = \eta_2 \times \frac{(\eta_3 + \eta_2)(\cos \beta_2 z - j \sin \beta_2 z) + (\eta_3 - \eta_2)(\cos \beta_2 z + j \sin \beta_2 z)}{(\eta_3 + \eta_2)(\cos \beta_2 z - j \sin \beta_2 z) - (\eta_3 - \eta_2)(\cos \beta_2 z + j \sin \beta_2 z)}$$

This is easily simplified to yield

$$\eta_w(z) = \eta_2 \frac{\eta_3 \cos \beta_2 z - j\eta_2 \sin \beta_2 z}{\eta_2 \cos \beta_2 z - j\eta_3 \sin \beta_2 z} \tag{35}$$

We now use the wave impedance in region 2 to solve our reflection problem. Of interest to us is the net reflected wave amplitude at the first interface. Since tangential **E** and **H** are continuous across the boundary, we have

$$E_{xs1}^+ + E_{xs1}^- = E_{xs2} \qquad (z = -l) \tag{36}$$

and

$$H_{ys1}^+ + H_{ys1}^- = H_{ys2} \qquad (z = -l) \tag{37}$$

Then, in analogy to (7) and (8), we may write

$$E_{x10}^+ + E_{x10}^- = E_{xs2}(z = -l) \tag{38}$$

and

$$\frac{E_{x10}^{+}}{\eta_1} - \frac{E_{x10}^{-}}{\eta_1} = \frac{E_{xs2}(z=-l)}{\eta_w(-l)} \tag{39}$$

where E_{x10}^{+} and E_{x10}^{-} are the amplitudes of the incident and reflected fields. We call $\eta_w(-l)$ the *input impedance*, η_{in}, to the two-interface combination. We now solve (38) and (39) together, eliminating E_{xs2}, to obtain

$$\boxed{\frac{E_{x10}^{-}}{E_{x10}^{+}} = \Gamma = \frac{\eta_{in} - \eta_1}{\eta_{in} + \eta_1}} \tag{40}$$

To find the input impedance, we evaluate (35) at $z = -l$, resulting in

$$\boxed{\eta_{in} = \eta_2 \frac{\eta_3 \cos \beta_2 l + j\eta_2 \sin \beta_2 l}{\eta_2 \cos \beta_2 l + j\eta_3 \sin \beta_2 l}} \tag{41}$$

Equations (40) and (41) are general results that enable us to calculate the net reflected wave amplitude and phase from two parallel interfaces between lossless media.[1] Note the dependence on the interface spacing, l, and on the wavelength as measured in region 2, characterized by β_2. Of immediate importance to us is the fraction of the incident power that reflects from the dual interface and back-propagates in region 1. As we found earlier, this fraction will be $|\Gamma|^2$. Also of interest is the transmitted power, which propagates away from the second interface in region 3. It is simply the remaining power fraction, which is $1 - |\Gamma|^2$. The power in region 2 stays constant in steady state; power leaves that region to form the reflected and transmitted waves, but is immediately replenished by the incident wave.

An important result of situations involving two interfaces is that it is possible to achieve total transmission in certain cases. From (40), we see that total transmission occurs when $\Gamma = 0$, or when $\eta_{in} = \eta_1$. In this case we say that the input impedance is *matched* to that of the incident medium. There are a few methods of accomplishing this.

As a start, suppose that $\eta_3 = \eta_1$, and region 2 is of thickness such that $\beta_2 l = m\pi$, where m is an integer. Now $\beta_2 = 2\pi/\lambda_2$, where λ_2 is the wavelength *as measured in region 2*. Therefore

$$\frac{2\pi}{\lambda_2} l = m\pi$$

[1] For convenience, (38) and (39) have been written for a specific time at which the incident wave amplitude, E_{x10}^{+}, occurs at $z = -l$. This establishes a zero-phase reference at the front interface for the incident wave, and so it is from this reference that the reflected wave phase is determined. Equivalently, we have repositioned the $z = 0$ point at the front interface. Eq. (41) allows this, since it is only a function of the interface spacing, l.

or

$$\boxed{l = m\frac{\lambda_2}{2}} \tag{42}$$

With $\beta_2 l = m\pi$, the second region thickness is an integer multiple of half-wavelengths as measured in that medium. Equation (41) now reduces to $\eta_{in} = \eta_3$. *Thus the general effect of a multiple half-wave thickness is to render the second region immaterial to the results on reflection and transmission*. Equivalently, we have a single interface problem involving η_1 and η_3. Now, with $\eta_3 = \eta_1$, we have a matched input impedance, and there is no net reflected wave. This method of choosing the region 2 thickness is known as *half-wave matching*. Its applications include, for example, antenna housings on airplanes known as *radomes*, which form a part of the fuselage. The antenna, inside the aircraft, can transmit and receive through this layer which can be shaped to enable good aerodynamic characteristics. Note that the half-wave matching condition no longer applies as we deviate from the wavelength that satisfies it. When this is done, the device reflectivity increases (with increased wavelength deviation), so it ultimately acts as a bandpass filter.

Another application, typically seen in optics, is the *Fabry-Perot interferometer*. This, in its simplest form, consists of a single block of glass or other transparent material, whose thickness, l, is set to transmit wavelengths which satisfy the condition, $\lambda_2 = 2l/m$. Often, it is desired to transmit only one wavelength, not several, as (42) would allow. We would therefore like to ensure that adjacent wavelengths that are passed through the device are separated as far as possible. This separation is in general given by

$$\lambda_{m-1} - \lambda_m = \Delta\lambda_f = \frac{2l}{m-1} - \frac{2l}{m} = \frac{2l}{m(m-1)} \doteq \frac{2l}{m^2}$$

Note that m is the number of half-wavelengths in region 2, or $m = 2l/\lambda_2$, where λ_2 is the desired wavelength for transmission. Thus

$$\Delta\lambda_f \doteq \frac{\lambda_2^2}{2l} \tag{43}$$

$\Delta\lambda_f$ is known as the *free spectral range* of the Fabry-Perot interferometer. The interferometer can be used as a narrow-band filter (transmitting a desired wavelength and a narrow spectrum around this wavelength) if the spectrum to be filtered is narrower than the free spectral range.

Example 12.4

Suppose we wish to filter an optical spectrum of full-width $\Delta\lambda_s = 50$ nm, and whose center wavelength is in the red part of the visible spectrum at 600 nm, where one nm (nanometer) is 10^{-9} m. A Fabry-Perot filter is to be used, consisting of a lossless glass

plate in air, having relative permittivity $\epsilon'_R = \epsilon_R = 2.1$. We need to find the required range of glass thicknesses such that multiple wavelength orders will not be transmitted.

Solution. We require that the free spectral range be greater than the optical spectral width, or $\Delta\lambda_f > \Delta\lambda_s$. Thus, using (43)

$$l < \frac{\lambda_2^2}{2\Delta\lambda_s}$$

where

$$\lambda_2 = \frac{600}{\sqrt{2.1}} = 414\,\text{nm}$$

So

$$l < \frac{414^2}{2(50)} = 1.7 \times 10^3\,\text{nm} = 1.7\,\mu\text{m}$$

where 1μm (micrometer) $= 10^{-6}$ m. Fabricating a glass plate of this thickness or less is somewhat ridiculous to contemplate. Instead, what is often used is an air space of thickness on this order, between two thick plates whose surfaces on the sides opposite the air space are antireflection coated. This is in fact a more versatile configuration since the wavelength to be transmitted (and the free spectral range) can be adjusted by varying the plate separation.

Next we remove the restriction $\eta_1 = \eta_3$ and look for a way to produce zero reflection. Suppose we set $\beta_2 l = (2m-1)\pi/2$, or an odd multiple of $\pi/2$. This means that

$$\frac{2\pi}{\lambda_2} l = (2m-1)\frac{\pi}{2} \qquad (m = 1, 2, 3, \ldots)$$

or

$$\boxed{l = (2m-1)\frac{\lambda_2}{4}} \tag{44}$$

The thickness is an odd multiple of a quarter wavelength as measured in region 2. Under this condition (41) reduces to

$$\eta_{in} = \frac{\eta_2^2}{\eta_3} \tag{45}$$

Typically, we choose the second region impedance to allow matching between given impedances η_1 and η_3. To achieve total transmission, we require that $\eta_{in} = \eta_1$, so that the required second region impedance becomes

$$\boxed{\eta_2 = \sqrt{\eta_1 \eta_3}} \tag{46}$$

With the conditions given by (44) and (46) satisfied, we have performed *quarter-wave matching*. The design of anti reflective coatings for optical devices is based on this principle.

Example 12.5

We wish to coat a glass surface with an appropriate dielectric layer to provide total transmission from air to the glass at a wavelength of 570 nm. The glass has dielectric constant, $\epsilon_R = 2.1$. Determine the required dielectric constant for the coating and its minimum thickness.

Solution. The known impedances are $\eta_1 = 377\,\Omega$ and $\eta_3 = 377/\sqrt{2.1} = 260\,\Omega$. Using (46) we have

$$\eta_2 = \sqrt{(377)(260)} = 313\,\Omega$$

The dielectric constant of region 2 will then be

$$\epsilon_{R2} = \left(\frac{377}{313}\right)^2 = 1.45$$

The wavelength in region 2 will be

$$\lambda_2 = \frac{570}{\sqrt{1.45}} = 473\,\text{nm}$$

The minimum thickness of the dielectric layer is then

$$l = \frac{\lambda_2}{4} = 118\,\text{nm} = 0.118\,\mu\text{m}$$

The procedure in this section for evaluating wave reflection has involved calculating an effective impedance at the first interface, η_{in}, which is expressed in terms of the impedances that lie beyond the front surface. This process of *impedance transformation* is more apparent when we consider problems involving more than two interfaces.

For example, consider the three-interface situation shown in Fig. 12.7, where a wave is incident from the left in region 1. We wish to determine the fraction of the incident power that is reflected and back-propagates in region 1, and the fraction of the incident power that is transmitted into region 4. To do this, we need to find the input impedance at the front surface (the interface between regions 1 and 2). We start by transforming the impedance of region 4 to form the input impedance at the boundary between regions 2 and 3. This is shown as $\eta_{in,b}$ in the figure. Using (41), we have

$$\eta_{in,b} = \eta_3 \frac{\eta_4 \cos\beta_3 l_b + j\eta_3 \sin\beta_3 l_b}{\eta_3 \cos\beta_3 l_b + j\eta_4 \sin\beta_3 l_b} \tag{47}$$

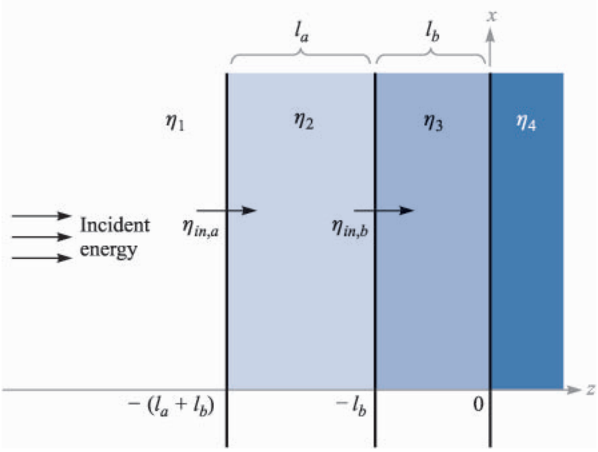


FIGURE 12.7
A three-interface problem, in which input impedance $\eta_{in,b}$ is transformed back to the front interface to form input impedance $\eta_{in,a}$.

We have now effectively reduced the situation to a two-interface problem in which $\eta_{in,b}$ is the impedance of all that lies beyond the second interface. The input impedance at the front interface, $\eta_{in,a}$, is now found by transforming $\eta_{in,b}$ as follows:

$$\eta_{in,a} = \eta_2 \frac{\eta_{in,b} \cos \beta_2 l_a + j\eta_2 \sin \beta_2 l_a}{\eta_2 \cos \beta_2 l_a + j\eta_{in,b} \sin \beta_2 l_a} \tag{48}$$

The reflected power fraction is now $|\Gamma|^2$, where

$$\Gamma = \frac{\eta_{in,a} - \eta_1}{\eta_{in,a} + \eta_1}$$

The fraction of the power transmitted into region 4 is, as before, $1 - |\Gamma|^2$. The method of impedance transformation can be applied in this manner to any number of interfaces. The process, although tedious, is easily handled by a computer.

The motivation for using multiple layers to reduce reflection is that the resulting structure is less sensitive to deviations from the design wavelength if the impedances are arranged to progressively increase or decrease from layer to layer. In using multiple layers to antireflection coat a camera lens, for example, the layer on the lens surface would be of impedance very close to that of the glass. Subsequent layers are given progressively higher impedances. With a large number of layers fabricated in this way, the situation begins to approach (but never reaches) the ideal case, in which the top layer impedance matches that of air, while the impedances of deeper layers continuously decrease until reaching the value of the glass surface. With this continuously varying impedance, there would be no surface from which to reflect, and so light of any wavelength is totally transmitted. Multilayer coatings designed in this way produce excellent broadband transmission characteristics.

The impedance transformation method for handling multiple interfaces applies not only to plane waves at boundaries, but also to loaded transmission lines of finite length, and to cascaded transmission lines. We will encounter problems of this type in the next chapter, which we will solve using exactly the same mathematics.

✔ **D12.3.** A uniform plane wave in air is normally incident on a dielectric slab of thickness $\lambda_2/4$, and intrinsic impedance $\eta_2 = 260\ \Omega$. Determine the magnitude and phase of the reflection coefficient.

Ans. 0.356; 180°.

12.4 PLANE WAVE PROPAGATION IN GENERAL DIRECTIONS

In this section we will learn how to mathematically describe uniform plane waves that propagate in any direction. Our motivation for doing this is our need to address the problem of incident waves on boundaries that are not perpendicular to the propagation direction. Such problems of *oblique incidence* generally occur, with normal incidence being a special case. Addressing such problems requires (as always) that we establish an appropriate coordinate system. With the boundary positioned in the x, y plane, for example, the incident wave will propagate in a direction that could involve all three coordinate axes, whereas with normal incidence, we were only concerned with propagation along z. We need a mathematical formalism that will allow for the general direction case.

Let us consider a wave that propagates in a lossless medium, with propagation constant $\beta = k = \omega\sqrt{\mu\epsilon}$. For simplicity, we consider a two-dimensional case, where the wave travels in a direction between the x and z axes. The first step is to consider the propagation constant as a *vector*, $\mathbf{k}$, indicated in Fig. 12.8. The direction of $\mathbf{k}$ is the propagation direction, which is the same as the direction of the Poynting vector in our case.[2] The magnitude of $\mathbf{k}$ is the phase shift per unit distance *along that direction*. Part of the process of characterizing a wave involves specifying its phase at any spatial location. For the waves we have considered that propagate along the z axis, this was accomplished by the factor $e^{\pm jkz}$ in the phasor form. To specify the phase in our two-dimensional problem, we make use of the vector nature of $\mathbf{k}$, and consider the phase at a general location, (x, z), described through the position vector $\mathbf{r}$. The phase at that location, referenced to the origin, is given by the projection of $\mathbf{k}$ along $\mathbf{r}$ times the magnitude of $\mathbf{r}$, or just

[2] We assume here that the wave is in an isotropic medium, where the permittivity and permeability do not change with field orientation. In anisotropic media (where ϵ and/or μ depend on field orientation), the directions of the Poynting vector and $\mathbf{k}$ may differ.

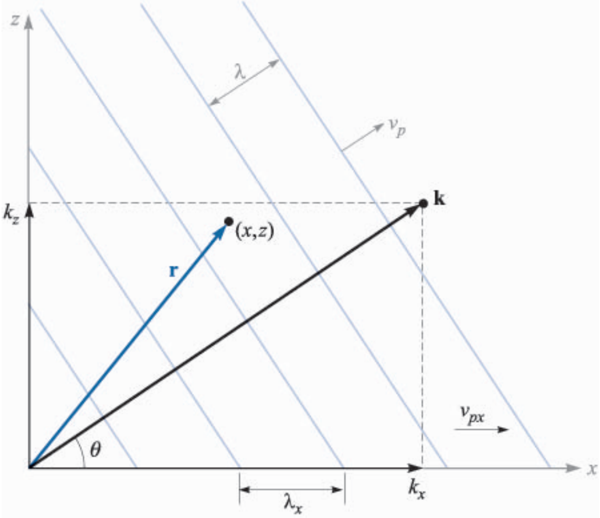


FIGURE 12.8 Representation of a uniform plane wave with wavevector **k** at angle θ to the x axis. The phase at point (x, z) is given by $\mathbf{k} \cdot \mathbf{r}$. Planes of constant phase (shown as lines perpendicular to **k**) are spaced by wavelength λ, but have wider spacing when measured along the x or z axes.

$\mathbf{k} \cdot \mathbf{r}$. If the electric field is of magnitude E_0, we can thus write down the phasor form of the wave in Fig. 12.8 as

$$\boxed{\mathbf{E}_s = \mathbf{E}_0 e^{-j\mathbf{k}\cdot\mathbf{r}}} \tag{49}$$

The minus sign in the exponent indicates that the phase along **r** moves in time in the direction of increasing **r**. Again, the wave power flow in an isotropic medium occurs in the direction along which the phase shift per unit distance is maximum—or along **k**. The vector **r** serves as a means to measure phase at any point using **k**. This construction is easily extended to three dimensions by allowing **k** and **r** to each have three components.

In our two-dimensional case of Fig. 12.8, we can express **k** in terms of its x and z components:

$$\boxed{\mathbf{k} = k_x \mathbf{a_x} + k_z \mathbf{a_z}}$$

The position vector, **r**, can be similarly expressed:

$$\boxed{\mathbf{r} = x\,\mathbf{a_x} + z\,\mathbf{a_z}}$$

so that

$$\mathbf{k} \cdot \mathbf{r} = k_x x + k_z z$$

Equation (49) now becomes

$$\mathbf{E}_s = \mathbf{E}_0 e^{-j(k_x x + k_z z)} \tag{50}$$

Whereas Eq. (49) provided the general form of the wave, Eq. (50) is the form that is specific to the situation. Given a wave expressed by (50), the angle of propagation from the x axis is readily found through

$$\theta = \tan^{-1}\left(\frac{k_z}{k_x}\right)$$

The wavelength and phase velocity depend on the direction one is considering. In the direction of $\mathbf{k}$, these will be

$$\lambda = \frac{2\pi}{k} = \frac{2\pi}{(k_x^2 + k_z^2)^{1/2}}$$

and

$$v_p = \frac{\omega}{k} = \frac{\omega}{(k_x^2 + k_z^2)^{1/2}}$$

If, for example, we consider the x direction, these quantities will be

$$\lambda_x = \frac{2\pi}{k_x}$$

and

$$v_{px} = \frac{\omega}{k_x}$$

Note that both λ_x and v_{px} are greater than their counterparts along the direction of $\mathbf{k}$. This result, at first surprising, can be understood through the geometry of Fig. 12.8. The diagram shows a series of phase fronts (planes of constant phase), which intersect $\mathbf{k}$ at right angles. The phase shift between adjacent fronts is set at 2π in the figure; this corresponds to a spatial separation along the $\mathbf{k}$ direction of one wavelength, as shown. The phase fronts intersect the x axis, and we see that *along* x the front separation is greater than it was along $\mathbf{k}$. λ_x is the spacing between fronts along x, and is indicated on the figure. The phase velocity along x is the velocity of the intersection points between the phase fronts and the x axis. Again, from the geometry, we see that this velocity must be faster than the velocity along $\mathbf{k}$, and will of course exceed the speed of light in the medium. This does not constitute a violation of special relativity, however, since the energy in the wave flows in the direction of $\mathbf{k}$, and not along x or z. The wave frequency is $f = \omega/2\pi$, and is invariant with direction. Note, for example, that in the directions we have considered,

$$f = \frac{v_p}{\lambda} = \frac{v_{px}}{\lambda_x} = \frac{\omega}{2\pi}$$

Example 12.6

Consider a 50 MHz uniform plane wave having electric field amplitude 10 V/m. The medium is lossless, having $\epsilon_R = \epsilon'_R = 9.0$ and $\mu_R = 1.0$. The wave propagates in the x, y plane at a 30° angle to the x axis, and is linearly polarized along z. Write down the phasor expression for the electric field.

Solution. The propagation constant magnitude is

$$k = \omega\sqrt{\mu\epsilon} = \frac{\omega\sqrt{\epsilon_R}}{c} = \frac{2\pi \times 50 \times 10^6(3)}{3 \times 10^8} = 3.14\,\mathrm{m}^{-1}$$

The vector **k** is now

$$\mathbf{k} = 3.14(\cos 30\,\mathbf{a_x} + \sin 30\,\mathbf{a_y}) = 2.7\mathbf{a_x} + 1.6\mathbf{a_y}\,\mathrm{m}^{-1}$$

Then

$$\mathbf{r} = x\,\mathbf{a_x} + y\,\mathbf{a_y}$$

With the electric field directed along z, the phasor form becomes

$$\mathbf{E_s} = E_0 e^{-j\mathbf{k}\cdot\mathbf{r}}\,\mathbf{a_z} = 10e^{-j(2.7x+1.6y)}\,\mathbf{a_z}$$

✔ **D12.4.** For Example 12.6, calculate λ_x, λ_y, v_{px}, and v_{py}.

Ans. 2.3 m; 3.9 m; 1.2×10^8 m/s; 2.0×10^8 m/s.

12.5 PLANE WAVE REFLECTION AT OBLIQUE INCIDENCE ANGLES

We now consider the problem of wave reflection from plane interfaces, in which the incident wave propagates at some angle to the surface. Our objectives are (1) to determine the relation between incident, reflected, and transmitted angles, and (2) to derive reflection and transmission coefficients that are functions of the incident angle and wave polarization. We will also show that cases exist in which total reflection or total transmission may occur at the interface between two dielectrics if the angle of incidence and the polarization are appropriately chosen.

The situation is illustrated in Fig. 12.9, in which the incident wave direction and position-dependent phase are characterized by wavevector, $\mathbf{k}_1^+$. The angle of incidence is the angle between $\mathbf{k}_1^+$ and a line that is normal to the surface (the x axis in this case). The incidence angle is shown as θ_1. The reflected wave, characterized by wavevector $\mathbf{k}_1^-$, will propagate away from the interface at angle θ'_1. Finally, the transmitted wave, characterized by $\mathbf{k}_2$, will propagate into the second region at angle θ_2 as shown. One would suspect (from previous experience) that the incident and reflected angles are equal ($\theta_1 = \theta'_1$), which is correct. We need to show this, however, to be complete.

The two media are lossless dielectrics, characterized by intrinsic impedances, η_1 and η_2. We will assume, as before, that the materials are non-

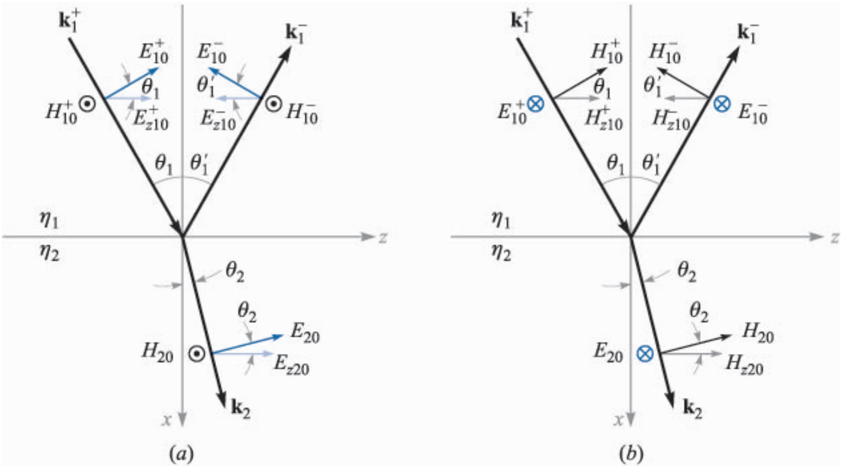


FIGURE 12.9
Geometries for plane wave incidence at angle θ_1 onto an interface between dielectrics having intrinsic impedances η_1 and η_2. The two polarization cases are shown: (*a*) p-polarization (or TM), with **E** in the plane of incidence; (*b*) s-polarization (or TE), with **E** perpendicular to the plane of incidence.

magnetic, and thus have permeability μ_0. Consequently, the materials are adequately described by specifying their dielectric constants, ϵ_{R1} and ϵ_{R2}. It is convenient at this stage to introduce the *refractive index* of the material, n, defined as the square root of the dielectric constant. We thus have $n_1 = \sqrt{\epsilon_{R1}}$ and $n_2 = \sqrt{\epsilon_{R2}}$. Use of either parameter (index or dielectric constant) is acceptable, with refractive index used almost exclusively in the optical frequency range (on the order of 10^{14} Hz); at lower frequencies, problems are typically posed using dielectric constants. It is important not to confuse index, n, with the similar-appearing Greek η (intrinsic impedance), which has an entirely different meaning.

In Fig. 12.9, two cases are shown which differ by the choice of electric field orientation. In Fig. 12.9*a*, the **E** field is polarized in the plane of the page, with **H** therefore perpendicular to the page and pointing outward. In this illustration, the plane of the page is also the *plane of incidence*, which is more precisely defined as the plane spanned by the incident **k** vector and the normal to the surface. With **E** lying in the plane of incidence, the wave is said to have *parallel polarization*, or is *p-polarized* (**E** is parallel to the incidence plane). Note that while **H** is perpendicular to the incidence plane, it lies parallel to the interface, or transverse to the direction normal to the interface. Consequently, another name for this type of polarization is *transverse magnetic*, or TM polarization.

Fig. 12.9*b* shows the situation in which the field directions have been rotated by 90°. Now **H** lies in the plane of incidence, whereas **E** is perpendicular to the plane. Since **E** is used to define polarization, the configuration is called

perpendicular polarization, or is *s-polarized*.[3] **E** is also parallel to the interface, and so the case is also called *transverse electric*, or TE polarization. We will find that the reflection and transmission coefficients will differ for the two polarization types, but that reflection and transmission angles will not depend on polarization. We only need to consider s and p polarizations, since any other field direction can be constructed as some combination of s and p waves.

Our desired knowledge of reflection and transmission coefficients, as well as how the angles relate, can be found through the field boundary conditions at the interface. Specifically, we require that the transverse components of **E** and **H** be continuous across the interface. These were the conditions we used to find Γ and τ for normal incidence ($\theta_1 = 0$), which is in fact a special case of our current problem. We will consider the case of p-polarization (Fig. 12.9*a*) first. To begin, we write down the incident, reflected, and transmitted fields in phasor form, using the notation developed in the previous section:

$$\mathbf{E}_{s1}^{+} = \mathbf{E}_{10}^{+}\, e^{-j\mathbf{k}_1^{+}\cdot\mathbf{r}} \tag{51}$$

$$\mathbf{E}_{s1}^{-} = \mathbf{E}_{10}^{-}\, e^{-j\mathbf{k}_1^{-}\cdot\mathbf{r}} \tag{52}$$

$$\mathbf{E}_{s2} = \mathbf{E}_{20}\, e^{-j\mathbf{k}_2\cdot\mathbf{r}} \tag{53}$$

where

$$\mathbf{k}_1^{+} = k_1(\cos\theta_1\,\mathbf{a_x} + \sin\theta_1\,\mathbf{a_z}) \tag{54}$$

$$\mathbf{k}_1^{-} = k_1\left(-\cos\theta_1'\,\mathbf{a_x} + \sin\theta_1'\,\mathbf{a_z}\right) \tag{55}$$

$$\mathbf{k}_2 = k_2(\cos\theta_2\,\mathbf{a_x} + \sin\theta_2\,\mathbf{a_z}) \tag{56}$$

and where

$$\mathbf{r} = x\,\mathbf{a_x} + z\,\mathbf{a_z} \tag{57}$$

The wavevector magnitudes are $k_1 = \omega\sqrt{\epsilon_{R1}}/c = n_1\omega/c$ and $k_2 = \omega\sqrt{\epsilon_{R2}}/c = n_2\omega/c$.

Now, to evaluate the boundary condition that requires continuous tangential electric field, we need to find the components of the electric fields (z components) that are parallel to the interface. Projecting all **E** fields in the z direction, and using (51) through (57), we find

$$E_{zs1}^{+} = E_{z10}^{+}\, e^{-j\mathbf{k}_1^{+}\cdot\mathbf{r}} = E_{10}^{+}\cos\theta_1 e^{-jk_1(x\cos\theta_1 + z\sin\theta_1)} \tag{58}$$

$$E_{zs1}^{-} = E_{z10}^{-}\, e^{-j\mathbf{k}_1^{-}\cdot\mathbf{r}} = E_{10}^{-}\cos\theta_1' e^{jk_1(x\cos\theta_1' - z\sin\theta_1')} \tag{59}$$

$$E_{zs2} = E_{z20}\, e^{-j\mathbf{k}_2\cdot\mathbf{r}} = E_{20}\cos\theta_2 e^{-jk_2(x\cos\theta_2 + z\sin\theta_2)} \tag{60}$$

[3] The *s* designation is an abbreviation for the German *senkrecht*, meaning *perpendicular*. The *p* in p-polarized is an abbreviation for the German word for parallel, which is *parallel*.

The boundary condition for continuous tangential electric field now reads:

$$E^{+}_{zs1} + E^{-}_{zs1} = E_{zs2} \quad (\text{at } x = 0)$$

We now substitute Eqs. (58) through (60) into (61), and evaluate the result at $x = 0$ to obtain:

$$E^{+}_{10} \cos\theta_1 \, e^{-jk_1 z \sin\theta_1} + E^{-}_{10} \cos\theta'_1 \, e^{-jk_1 z \sin\theta'_1} = E_{20} \cos\theta_2 \, e^{-jk_2 z \sin\theta_2} \tag{61}$$

Note that E^{+}_{10}, E^{-}_{10}, and E_{20} are all constants (independent of z). Further, we require that (61) hold for all values of z (everywhere on the interface). For this to occur, it must follow that all the phase terms appearing in (61) are equal. Specifically,

$$k_1 z \sin\theta_1 = k_1 z \sin\theta'_1 = k_2 z \sin\theta_2$$

From this, we see immediately that $\theta'_1 = \theta_1$, or the angle of reflection is equal to the angle of incidence. We also find that

$$\boxed{k_1 \sin\theta_1 = k_2 \sin\theta_2} \tag{62}$$

Eq. (62) is known as *Snell's law of refraction*. Since in general, $k = n\omega/c$, we can rewrite (62) in terms of the refractive indices:

$$\boxed{n_1 \sin\theta_1 = n_2 \sin\theta_2} \tag{63}$$

Eq. (63) is the form of Snell's law that is most readily used for our present case of nonmagnetic dielectrics. Eq. (62) is a more general form which would apply, for example, to cases involving materials with different permeabilities as well as different permittivities. In general, we would have $k_1 = (\omega/c)\sqrt{\mu_{R1}\epsilon_{R1}}$ and $k_2 = (\omega/c)\sqrt{\mu_{R2}\epsilon_{R2}}$.

Having found the relations between angles, we next turn to our second objective, which is to determine the relations between the amplitudes, E^{+}_{10}, E^{-}_{10}, and E_{20}. To accomplish this, we need to consider the other boundary condition, requiring tangential continuity of **H** at $x = 0$. The magnetic field vectors for the p-polarized wave are all negative y-directed. At the boundary the field amplitudes are related through

$$H^{+}_{10} + H^{-}_{10} = H_{20} \tag{64}$$

Then, using the fact that $\theta'_1 = \theta_1$, and invoking Snell's law, (61) becomes

$$E^{+}_{10} \cos\theta_1 + E^{-}_{10} \cos\theta_1 = E_{20} \cos\theta_2 \tag{65}$$

Using the medium intrinsic impedances, we know for example, that $E^{+}_{10}/H^{+}_{10} = \eta_1$ and $E^{+}_{20}/H^{+}_{20} = \eta_2$. Eq. (64) can be written as follows:

$$\frac{E_{10}^{+}\cos\theta_1}{\eta_{1p}} - \frac{E_{10}^{-}\cos\theta_1}{\eta_{1p}} = \frac{E_{20}^{+}\cos\theta_2}{\eta_{2p}} \tag{66}$$

Note the minus sign in front of the second term in (66), which results from the fact that $E_{10}^{-}\cos\theta_1$ is negative (from Fig. 12.9*a*), whereas H_{10}^{-} is positive (again from the figure). In writing Eq. (66), *effective impedances*, valid for p-polarization, are defined through

$$\boxed{\eta_{1p} = \eta_1 \cos\theta_1} \tag{67}$$

and

$$\boxed{\eta_{2p} = \eta_2 \cos\theta_2} \tag{68}$$

Using this representation, Eqs. (65) and (66) are now in a form that enables them to be solved together for the ratios E_{10}^{-}/E_{10}^{+} and E_{20}/E_{10}^{+}. Performing analogous procedures to those used in solving (7) and (8), we find the reflection and transmission coefficients:

$$\boxed{\Gamma_p = \frac{E_{10}^{-}}{E_{10}^{+}} = \frac{\eta_{2p} - \eta_{1p}}{\eta_{2p} + \eta_{1p}}} \tag{69}$$

$$\boxed{\tau_p = \frac{E_{20}}{E_{10}^{+}} = \frac{2\eta_{2p}}{\eta_{2p} + \eta_{1p}}} \tag{70}$$

A similar procedure can be carried out for s-polarization, referring to Fig. 12.9*b*. The details are left as an exercise; the results are

$$\boxed{\Gamma_s = \frac{E_{y10}^{-}}{E_{y10}^{+}} = \frac{\eta_{2s} - \eta_{1s}}{\eta_{2s} + \eta_{1s}}} \tag{71}$$

$$\boxed{\tau_s = \frac{E_{y20}}{E_{y10}^{+}} = \frac{2\eta_{2s}}{\eta_{2s} + \eta_{1s}}} \tag{72}$$

where the effective impedances for s-polarization are

$$\boxed{\eta_{1s} = \eta_1 \sec\theta_1} \tag{73}$$

and

$$\boxed{\eta_{2s} = \eta_2 \sec\theta_2} \tag{74}$$

Equations (67) through (74) are what we need to calculate wave reflection and transmission for either polarization, and at any incident angle.

Example 12.7

A uniform plane wave is incident from air onto glass at an angle from the normal of 30°. Determine the fraction of the incident power that is reflected and transmitted for (a) p-polarization and (b) s-polarization. Glass has refractive index $n_2 = 1.45$.

Solution. First, we apply Snell's law to find the transmission angle. Using $n_1 = 1$ for air, we use (63) to find

$$\theta_2 = \sin^{-1}\left(\frac{\sin 30}{1.45}\right) = 20.2^\circ$$

Now, for p-polarization:

$$\eta_{1p} = \eta_1 \cos 30 = (377)(.866) = 326\,\Omega$$

$$\eta_{2p} = \eta_2 \cos 20.2 = \frac{377}{1.45}(.938) = 244\,\Omega$$

Then, using (69), we find

$$\Gamma_p = \frac{244 - 326}{244 + 326} = -0.144$$

The fraction of the incident power that is reflected is

$$\frac{P_r}{P_{inc}} = |\Gamma_p|^2 = .021$$

The transmitted fraction is then

$$\frac{P_t}{P_{inc}} = 1 - |\Gamma_p|^2 = .979$$

For s-polarization, we have

$$\eta_{1s} = \eta_1 \sec 30 = 377/.866 = 435\,\Omega$$

$$\eta_{2s} = \eta_2 \sec 20.2 = \frac{377}{1.45(.938)} = 277\,\Omega$$

Then, using (71):

$$\Gamma_s = \frac{277 - 435}{277 + 435} = -.222$$

The reflected power fraction is thus

$$|\Gamma_s|^2 = .049$$

The fraction of the incident power that is transmitted is

$$1 - |\Gamma_s|^2 = .951$$

In the previous example, reflection coefficient values for the two polarizations were found to be negative. The meaning of a negative reflection coefficient is that the component of the reflected electric field that is parallel to the interface will be directed opposite the incident field component when both are evaluated at the boundary.

The above effect is also observed when the second medium is a perfect conductor. In this case, we know that the electric field inside the conductor must be zero. Consequently, $\eta_2 = E_{20}/H_{20} = 0$, and the reflection coefficients will be $\Gamma_p = \Gamma_s = -1$. Total reflection occurs, regardless of the incident angle or polarization.

Now that we have methods available to us for solving problems involving oblique incidence reflection and transmission, we can explore the special cases of *total reflection* and *total transmission*. We look for special combinations of media, incidence angles, and polarizations that produce these properties. To begin, we identify the necessary condition for total reflection. We want total *power* reflection, so that $|\Gamma|^2 = \Gamma\Gamma^* = 1$, where Γ is either Γ_p or Γ_s. The fact that this condition involves the possibility of a complex Γ allows some flexibility. For the incident medium, we note that η_{1p} and η_{1s} will always be real and positive. On the other hand, when we consider the second medium, η_{2p} and η_{2s} involve factors of $\cos\theta_2$ or $1/\cos\theta_2$, where

$$\cos\theta_2 = \left[1 - \sin^2\theta_2\right]^{1/2} = \left[1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1\right]^{1/2} \tag{75}$$

where Snell's law has been used. We observe that $\cos\theta_2$, and hence η_{2p} and η_{2s}, become imaginary whenever $\sin\theta_1 > n_2/n_1$. Let us consider parallel polarization for example. Under conditions of imaginary η_{2p}, (69) becomes

$$\Gamma_p = \frac{j|\eta_{2p}| - \eta_{1p}}{j|\eta_{2p}| + \eta_{1p}} = -\frac{\eta_{1p} - j|\eta_{2p}|}{\eta_{1p} + j|\eta_{2p}|} = -\frac{Z}{Z^*}$$

where $Z = \eta_{1p} - j|\eta_{2p}|$. *We can therefore see that* $\Gamma_p\Gamma_p^* = 1$, *meaning total power reflection, whenever* η_{2p} *is imaginary*. The same will be true whenever η_{2p} is zero, which will occur when $\sin\theta_1 = n_2/n_1$. We thus have our condition for total reflection, which is

$$\boxed{\sin\theta_1 \geq \frac{n_2}{n_1}} \tag{76}$$

From this condition arises the *critical angle* of total reflection, θ_c, defined through

$$\boxed{\sin\theta_c = \frac{n_2}{n_1}} \tag{77}$$

The total reflection condition can thus be more succinctly written as

$$\boxed{\theta_1 \geq \theta_c \qquad \text{(for total reflection)}} \tag{78}$$

Note that for (76) and (77) to make sense, it must be true that $n_2 < n_1$, or the wave must be incident from a medium of higher refractive index than that of the medium beyond the boundary. For this reason, the total reflection condition is sometimes called total *internal* reflection; it is often seen (and applied) in optical devices such as beam-steering prisms, where light within the glass structure totally reflects from glass-air interfaces.

Example 12.8

A prism is to be used to turn a beam of light by 90°, as shown in Fig. 12.10. Light enters and exits the prism through two antireflective (AR-coated) surfaces. Total reflection is to occur at the back surface, where the incident angle is 45° to the normal. Determine the minimum required refractive index of the prism material if the surrounding region is air.

Solution. Considering the back surface, the medium beyond the interface is air, with $n_2 = 1.00$. Since $\theta_1 = 45°$, (76) is used to obtain

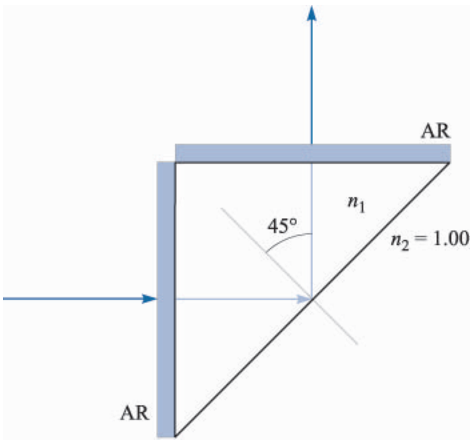


FIGURE 12.10
Beam-steering prism for Example 12.8.

$$n_1 \geq \frac{n_2}{\sin 45} = \sqrt{2} = 1.41$$

Since fused silica glass has refractive index $n_g = 1.45$, it is a suitable material for this application, and is in fact widely used.

Another important application of total reflection is in *optical waveguides*. These, in their simplest form, are constructed of three layers of glass, in which the middle layer has a slightly higher refractive index than the outer two. Fig. 12.11 shows the basic structure. Light, propagating from left to right, is confined to the middle layer by total reflection at the two interfaces, as shown. Optical fiber waveguides are constructed on this principle, in which a cylindrical glass core region of small radius is surrounded coaxially by a lower-index cladding glass material of larger radius. Basic waveguiding principles as applied to metallic and dielectric structures will be presented in Chapter 14.

We next consider the possibility of *total transmission*. In this case the requirement is simply that $\Gamma = 0$. We investigate this possibility for the two polarizations. First, we consider s-polarization. If $\Gamma_s = 0$, then from (71) we require that $\eta_{2s} = \eta_{1s}$, or

$$\eta_2 \sec\theta_2 = \eta_1 \sec\theta_1$$

Using Snell's law to write θ_2 in terms of θ_1, the above becomes

$$\eta_2\left[1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1\right]^{-1/2} = \eta_1\left[1 - \sin^2\theta_1\right]^{-1/2}$$

There is no value of θ_1 that will satisfy the above, so we turn instead to p-polarization. Using (67), (68), and (69), with Snell's law, the condition for $\Gamma_p = 0$ is

$$\eta_2\left[1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1\right]^{1/2} = \eta_1\left[1 - \sin^2\theta_1\right]^{1/2}$$

This equation does have a solution, which is

$$\boxed{\sin\theta_1 = \sin\theta_B = \frac{n_2}{\sqrt{n_1^2 + n_2^2}}} \tag{79}$$

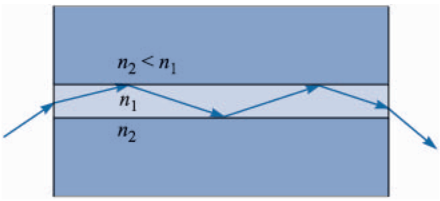


FIGURE 12.11
A dielectric slab waveguide (symmetric case), showing light confinement to the center material by total reflection.

where we have used $\eta_1 = \eta_0/n_1$ and $\eta_2 = \eta_0/n_2$. We call this special angle, θ_B, where total transmission occurs, the *Brewster angle* or *polarization angle*. The latter name comes from the fact that if light having both s- and p-polarization components is incident at $\theta_1 = \theta_B$, the p component will be totally transmitted, leaving the partially reflected light entirely s-polarized. At angles that are slightly off the Brewster angle, the reflected light is still predominantly s-polarized. Most reflected light that we see originates from horizontal surfaces (such as the surface of the ocean), and as such, the light is mostly of horizontal polarization. Polaroid sunglasses take advantage of this fact to reduce glare, since they are made to block transmission of horizontally polarized light, while passing light that is vertically polarized.

Example 12.9

Light is incident from air to glass at Brewster's angle. Determine the incident and transmitted angles.

Solution. Since glass has refractive index $n_2 = 1.45$, the incident angle will be

$$\theta_1 = \theta_B = \sin^{-1}\left(\frac{n_2}{\sqrt{n_1^2 + n_2^2}}\right) = \sin^{-1}\left(\frac{1.45}{\sqrt{1.45^2 + 1}}\right) = 55.4^\circ$$

The transmitted angle is found from Snell's law, through

$$\theta_2 = \sin^{-1}\left(\frac{n_1}{n_2}\sin\theta_B\right) = \sin^{-1}\left(\frac{n_1}{\sqrt{n_1^2 + n_2^2}}\right) = 34.6^\circ$$

Note from this exercise that $\sin\theta_2 = \cos\theta_B$, which means that the sum of the incident and refracted angles at the Brewster condition is always 90°.

✔ **D12.5.** In Example 12.9, calculate the reflection coefficient for s-polarized light.

Ans. -0.355

Many of the results we have seen in this section are summarized in Fig. 12.12, in which Γ_p and Γ_s, from (69) and (71), are plotted as functions of the incident angle, θ_1. Curves are shown for selected values of the refractive index ratio, n_1/n_2. For all plots in which $n_1/n_2 > 1$, Γ_s and Γ_p achieve a value of ± 1 at the critical angle. At larger angles, the reflection coefficients become imaginary (and are not shown) but nevertheless retain magnitudes of unity. The occurrence of the Brewster angle is evident in the curves for Γ_p (Fig. 12.12*a*), as all curves cross the θ_1 axis. This behavior is not seen in the Γ_s functions (Fig. 12.12*b*), as Γ_s is positive for all values of θ_1 when $n_1/n_2 > 1$, and is negative for $n_1/n_2 < 1$.

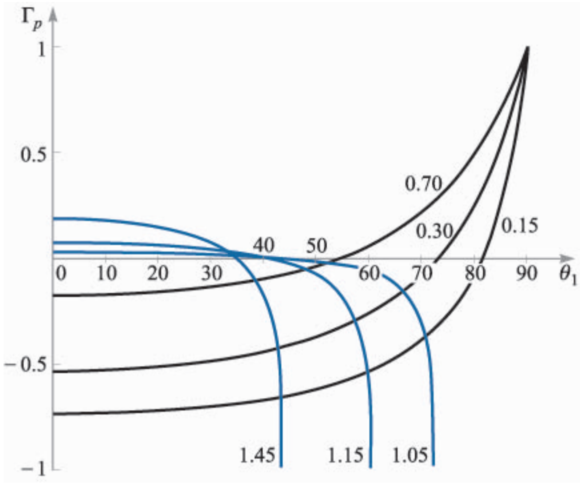


FIGURE 12.12*a*
Plots of Γ_p (Eq. (69)) as functions of the incident angle, θ_1, as per Fig. 12.9*a*. Curves are shown for selected values of the refractive index ratio, n_1/n_2. Both media are lossless and have $\mu_R = 1$. Thus $\eta_1 = \eta_0/n_1$ and $\eta_2 = \eta_0/n_2$.

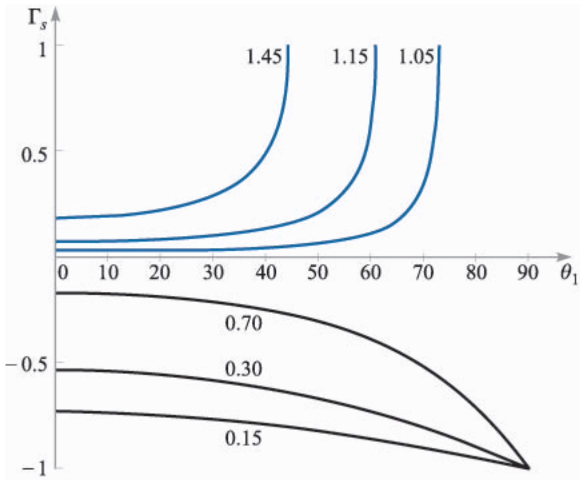


FIGURE 12.12*b*
Plots of Γ_s (Eq. (71)) as functions of the incident angle, θ_1, as per Fig. 12.9*b*. As in Fig. 12.12*a*, the media are lossless, and curves are shown for selected n_1/n_2.

12.6 WAVE PROPAGATION IN DISPERSIVE MEDIA

In Chapter 11, we encountered situations in which the complex permittivity of the medium depends on frequency. This is true in all materials through a number of possible mechanisms. One of these, mentioned earlier, is that oscillating bound charges in a material are in fact harmonic oscillators that have resonant frequencies associated with them (see Appendix D). When the frequency of an incoming electromagnetic wave is at or near a bound charge resonance, the wave will induce strong oscillations; these in turn have the effect of depleting energy from the wave in its original form. The wave thus experiences absorption, and

does so to a greater extent than it would at a frequency that is detuned from resonance. A related effect is that the real part of the dielectric constant will be different at frequencies near resonance than at frequencies far from resonance. In short, resonance effects give rise to values of ϵ' and ϵ'' that will vary continuously with frequency. These in turn will produce a fairly complicated frequency dependence in the attenuation and propagation constants as expressed in Eqs. (35) and (36) in Chapter 11.

This section concerns the effect of a frequency-varying dielectric constant (or refractive index) on a wave as it propagates in an otherwise lossless medium. This situation in fact arises quite often, since significant refractive index variation can occur at frequencies far away from resonance, where absorptive losses are negligible. A classic example of this is the separation of white light into its component colors by a glass prism. In this case, the frequency-dependent refractive index results in different angles of refraction for the different colors—so hence the separation. The color separation effect produced by the prism is known as *angular dispersion*, or more specifically, *chromatic* angular dispersion.

The term *dispersion* implies a *separation* of distinguishable components of a wave. In the case of the prism, the components are the various colors that have been spatially separated. An important point here is that the spectral *power* has been dispersed by the prism. We can motivate this idea by considering what it would take to measure the difference in refracted angles between, for example, blue and red light. One would need to use a power detector with a very narrow aperture, as shown in Fig. 12.13. The detector would be positioned at the locations of the blue and red light from the prism, with the narrow aperture allowing essentially one color at a time (or light over a very narrow spectral range) to pass through to the detector. The detector would thus measure the power in what we could call a "spectral packet," or a very narrow slice of the total power spectrum. The smaller the aperture, the narrower the spectral width of the packet, and thus the greater the precision in the measurement.[4] It is important for us to think of wave power as subdivided into spectral packets in this way, as it will figure

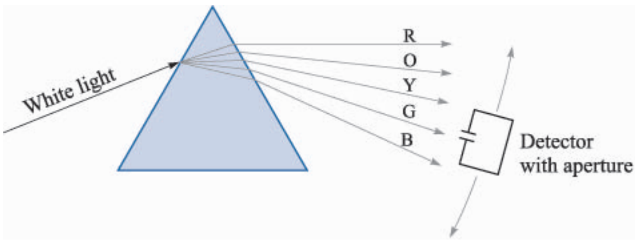


FIGURE 12.13
The angular dispersion of a prism can be measured using a movable device which measures both wavelength and power. The device senses light through a small aperture, thus improving wavelength resolution.

[4] To perform this experiment, one would need to measure the wavelength as well. Thus, the detector would likely be located at the output of a spectrometer or monochrometer, whose input slit performs the function of the bandwidth-limiting aperture.

prominently in our interpretation of the main topic of this section, which is wave dispersion *in time*.

We now consider a lossless nonmagnetic medium in which the refractive index varies with frequency. The propagation constant of a uniform plane wave in this medium will assume the form:

$$\beta(\omega) = k = \omega\sqrt{\mu_0 \epsilon(\omega)} = n(\omega)\frac{\omega}{c} \tag{80}$$

If we take $n(\omega)$ to be a monotonically increasing function of frequency (as is usually the case), a plot of ω vs. β would look something like the curve shown in Fig. 12.14. Such a plot is known as an ω-β *diagram* for the medium. Much can be learned about how waves propagate in the material by considering the shape of the ω-β curve.

Suppose we have two waves at two frequencies, ω_a and ω_b, which are co-propagating in the material, and whose amplitudes are equal. The two frequencies are labeled on the curve in Fig. 12.14, along with the frequency mid-way between the two, ω_0. The corresponding propagation constants, β_a, β_b, and β_0 are also labeled. The electric fields of the two waves are linearly polarized in the same direction (along x for example), while both waves propagate in the forward z direction. The waves will thus interfere with each other, producing a resultant wave whose field function can be found simply by adding the **E** fields of the two waves. This addition is done using the complex fields:

$$E_{c,net}(z,t) = E_0\left[e^{-j\beta_a z}e^{j\omega_a t} + e^{-j\beta_b z}e^{j\omega_b t}\right]$$

Note that we must use the full complex forms (with frequency dependence retained) as opposed to the phasor forms, since the waves are at different frequencies. Next, we factor out the term $e^{-j\beta_0 z}e^{j\omega_0 t}$:

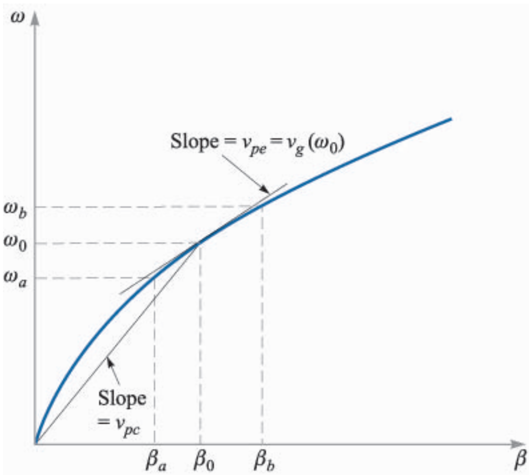


FIGURE 12.14
ω-β diagram for a material in which refractive index increases with frequency. The slope of a line tangent to the curve at ω_0 is the group velocity at that frequency. The slope of a line joining the origin to the point on the curve at ω_0 is the phase velocity at ω_0.

$$E_{c,net}(z,t) = E_0 e^{-j\beta_0 z} e^{j\omega_0 t}\left[e^{j\Delta\beta z} e^{-j\Delta\omega t} + e^{-j\Delta\beta z} e^{j\Delta\omega t}\right] = 2E_0 e^{-j\beta_0 z} e^{j\omega_0 t}\cos(\Delta\omega t - \Delta\beta z) \tag{81}$$

where

$$\Delta\omega = \omega_0 - \omega_a = \omega_b - \omega_0$$

and

$$\Delta\beta = \beta_0 - \beta_a = \beta_b - \beta_0$$

The above expression for $\Delta\beta$ is approximately true as long as $\Delta\omega$ is small. This can be seen from Fig. 12.14, by observing how the shape of the curve affects $\Delta\beta$, given uniform frequency spacings.

The real instantaneous form of (81) is found through

$$E_{net}(z,t) = \text{Re}\{E_{c,net}\} = 2E_0 \cos(\Delta\omega t - \Delta\beta z)\cos(\omega_0 t - \beta_0 z) \tag{82}$$

If $\Delta\omega$ is fairly small compared to ω_0, we recognize (82) as a carrier wave at frequency ω_0 that is sinusoidally modulated at frequency $\Delta\omega$. The two original waves are thus "beating" together to form a slow modulation, as one would hear when the same note is played by two slightly out-of-tune musical instruments. The resultant wave is shown in Fig. 12.15.

Of interest to us are the phase velocities of the carrier wave and the modulation envelope. From (82), we can immediately write these down as:

$$v_{pc} = \frac{\omega_0}{\beta_0} \qquad \text{(carrier velocity)} \tag{83}$$

$$v_{pe} = \frac{\Delta\omega}{\Delta\beta} \qquad \text{(envelope velocity)} \tag{84}$$

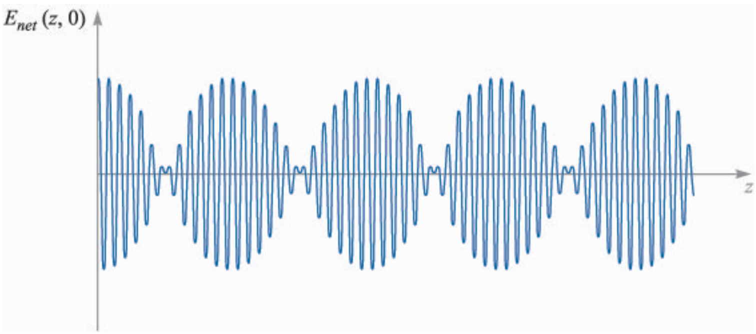


FIGURE 12.15
Plot of the total electric field strength as a function of z (with $t = 0$) of two co-propagating waves having different frequencies, ω_a and ω_b, as per Eq. (82). The rapid oscillations are associated with the carrier frequency, $\omega_0 = (\omega_a + \omega_b)/2$. The slower modulation is associated with the envelope or "beat" frequency, $\Delta\omega = (\omega_b - \omega_a)/2$.

Referring to the ω-β diagram, Fig. 12.14, we recognize the carrier phase velocity as the slope of the straight line that joins the origin to the point on the curve whose coordinates are ω_0 and β_0. We recognize the envelope velocity as a quantity that approximates the slope of the ω-β curve at the location of an operation point specified by (ω_0, β_0). The envelope velocity in this case is thus somewhat less than the carrier velocity. As $\Delta\omega$ becomes vanishingly small, the envelope velocity is identically the slope of the curve at ω_0. We can thus state the following for our example:

$$\lim_{\Delta\omega\to 0} \frac{\Delta\omega}{\Delta\beta} = \left.\frac{d\omega}{d\beta}\right|_{\omega_0} = v_g(\omega_0) \tag{85}$$

The quantity $d\omega/d\beta$ is called the *group velocity* function for the material, $v_g(\omega)$. When evaluated at a specified frequency, ω_0, it represents the velocity of a group of frequencies within a spectral packet of vanishingly small width, centered at frequency ω_0. In stating this, we have extended our two-frequency example to include waves that have a continuous frequency spectrum. To each frequency component (or packet) is associated a group velocity at which the energy in that packet propagates. Since the slope of the ω-β curve changes with frequency, group velocity will obviously be a function of frequency. The *group velocity dispersion* of the medium is, to first order, the rate at which the slope of the ω-β curve changes with frequency. It is this behavior that is of critical practical importance to the propagation of modulated waves within dispersive media, and the extent to which the modulation envelope may degrade with propagation distance.

Example 12.10

Consider a medium in which the refractive index varies linearly with frequency over a certain range:

$$n(\omega) = n_0 \frac{\omega}{\omega_0}$$

Determine the group velocity and the phase velocity of a wave at frequency ω_0.

Solution. First, the phase constant will be

$$\beta(\omega) = n(\omega)\frac{\omega}{c} = \frac{n_0\omega^2}{\omega_0 c}$$

Now

$$\frac{d\beta}{d\omega} = \frac{2n_0\omega}{\omega_0 c}$$

so that

$$v_g = \frac{d\omega}{d\beta} = \frac{\omega_0 c}{2n_0\omega}$$

The group velocity at ω_0 is

$$v_g(\omega_0) = \frac{c}{2n_0}$$

The phase velocity at ω_0 will be

$$v_p(\omega_0) = \frac{\omega}{\beta(\omega_0)} = \frac{c}{n_0}$$

To see how a dispersive medium affects a modulated wave, let us consider the propagation of an electromagnetic pulse. Pulses are used in digital signals, where the presence or absence of a pulse in a given time slot corresponds to a digital "one" or "zero." The effect of the dispersive medium on a pulse is to broaden it in time. To see how this happens, we consider the pulse *spectrum*, which is found through the Fourier transform of the pulse in time domain. In particular, suppose the pulse shape in time is Gaussian, and has electric field given at position $z = 0$ by

$$\boxed{E_c(0, t) = E_0 e^{-\frac{1}{2}(t/T)^2} e^{j\omega_0 t}} \tag{86}$$

where E_0 is a constant, ω_0 is the carrier frequency, and where T is the characteristic half-width of the pulse envelope; this is the time at which the pulse *intensity*, or magnitude of the Poynting vector, falls to $1/e$ of its maximum value (note that intensity is proportional to the square of the electric field). The frequency spectrum of the pulse is the Fourier transform of (86), which is

$$\boxed{E(0, \omega) = \frac{E_0 T}{\sqrt{2\pi}} e^{-\frac{1}{2}T^2(\omega-\omega_0)^2}} \tag{87}$$

Note from (87) that the frequency displacement from ω_0 at which the spectral *intensity* (proportional to $|E(0, \omega)|^2$) falls to $1/e$ of its maximum is $\Delta\omega = \omega - \omega_0 = 1/T$.

Fig. 12.16*a* shows the Gaussian intensity spectrum of the pulse, centered at ω_0, where the frequencies corresponding to the $1/e$ spectral intensity positions, ω_a and ω_b, are indicated. Fig. 12.16*b* shows the same three frequencies marked on the ω-β curve for the medium. Three lines are drawn that are tangent to the curve at the three frequency locations. The slopes of the lines indicate the group velocities at ω_a, ω_b, and ω_0, indicated as v_{ga}, v_{gb}, and v_{g0}. We can think of the pulse spreading in time as resulting from the differences in propagation times of the spectral energy packets that make up the pulse spectrum. Since the pulse spectral energy is highest at the center frequency, ω_0, we can use this as a reference point, about which further spreading of the energy will occur. For example, let us consider the difference in arrival times (group delays) between the frequency components, ω_0 and ω_b, after propagating through a distance z of the medium:

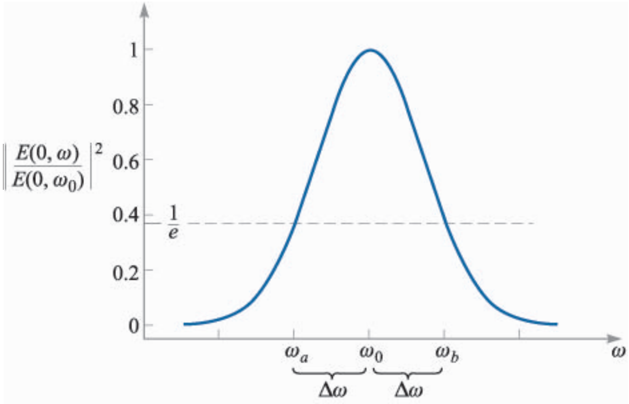


FIGURE 12.16*a*
Normalized power spectrum of a Gaussian pulse, as determined from Eq. (86). The spectrum is centered at carrier frequency ω_0, and has $1/e$ half-width, $\Delta\omega$. Frequencies ω_a and ω_b correspond to the $1/e$ positions on the spectrum.

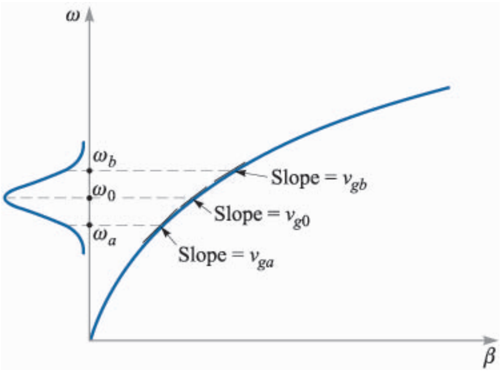


FIGURE 12.16*b*
The spectrum of Fig. 12.16*a* as shown on the ω-β diagram for the medium. The three frequencies specified in Fig. 12.16*a* are associated with three different slopes on the curve, resulting in different group delays for the spectral components.

$$\Delta\tau = z\left(\frac{1}{v_{gb}} - \frac{1}{v_{g0}}\right) = z\left(\left.\frac{d\beta}{d\omega}\right|_{\omega b} - \left.\frac{d\beta}{d\omega}\right|_{\omega 0}\right) \tag{88}$$

The gist of this result is that the medium is acting as what could be called a *temporal prism*. Instead of spreading out the spectral energy packets spatially, it is spreading them out in time. In so doing, a new temporal pulse envelope is constructed whose width is based fundamentally on the spread of propagation delays of the different spectral components. By determining the delay difference between the peak spectral component and the component at the spectral half-width, we construct an expression for the new *temporal* half-width. This assumes, of course, that the initial pulse width is negligible in comparison, but if not, we can account for that also, as will be shown later on.

To evaluate (88), we need more information about the ω-β curve. If we assume that the curve is smooth and has fairly uniform curvature, we can express $\beta(\omega)$ as the first three terms of a Taylor series expansion about the carrier frequency, ω_0:

$$\beta(\omega) \doteq \beta(\omega_0) + (\omega - \omega_0)\beta_1 + \frac{1}{2}(\omega - \omega_0)^2 \beta_2 \tag{89}$$

where

$$\beta_0 = \beta(\omega_0)$$

$$\beta_1 = \left.\frac{d\beta}{d\omega}\right|_{\omega 0} \tag{90}$$

and

$$\beta_2 = \left.\frac{d^2\beta}{d\omega^2}\right|_{\omega 0} \tag{91}$$

Note that if the ω-β curve were a straight line, then the first two terms in (89) would precisely describe $\beta(\omega)$. It is the third term in (89), involving β_2, that describes the curvature and ultimately the dispersion.

Noting that β_0, β_1, and β_2 are constants, we take the first derivative of (89) with respect to ω to find

$$\frac{d\beta}{d\omega} = \beta_1 + (\omega - \omega_0)\beta_2 \tag{92}$$

We now substitute (92) into (88) to obtain

$$\Delta\tau = [\beta_1 + (\omega_b - \omega_0)\beta_2]z - [\beta_1 + (\omega_0 - \omega_0)\beta_2]z = \Delta\omega\beta_2 z = \frac{\beta_2 z}{T} \tag{93}$$

where $\Delta\omega = (\omega_b - \omega_0) = 1/T$. β_2, as defined in Eq. (91), is the *dispersion parameter*. Its units are in general [time^2/distance], that is—pulse spread in time per unit spectral bandwidth, per unit distance. In optical fibers, for example, the units most commonly used are picoseconds^2/kilometer (psec^2/km). β_2 can be determined, knowing how β varies with frequency, or it can be measured.

If the initial pulse width is very short compared to $\Delta\tau$, then the broadened pulse width at location z will be simply $\Delta\tau$. If the initial pulse width is comparable to $\Delta\tau$, then the pulse width at z can be found through the convolution of the initial Gaussian pulse envelope of width T with a Gaussian envelope whose width is $\Delta\tau$. Thus in general, the pulse width at location z will be

$$T' = \sqrt{T^2 + (\Delta\tau)^2} \tag{94}$$

Example 12.11

An optical fiber channel is known to have dispersion, $\beta_2 = 20\ \text{ps}^2/\text{km}$. A Gaussian light pulse at the input of the fiber is of initial width $T = 10$ ps. Determine the width of the pulse at the fiber output, if the fiber is 15 km long.

Solution. The pulse spread will be

$$\Delta\tau = \frac{\beta_2 z}{T} = \frac{(20)(15)}{10} = 30\ \text{ps}$$

So the output pulse width is

$$T' = \sqrt{(10)^2 + (30)^2} = 32\ \text{ps}$$

An interesting by-product of pulse broadening through chromatic dispersion is that the broadened pulse is *chirped*. This means that the instantaneous frequency of the pulse varies monotonically (either increases or decreases) with time over the pulse envelope. This again is just a manifestation of the broadening mechanism, in which the spectral components at different frequencies are spread out in time as they propagate at different group velocities. We can quantify the effect by calculating the group delay, τ_g, as a function of frequency, using (92). We obtain:

$$\tau_g = \frac{z}{v_g} = z\frac{d\beta}{d\omega} = (\beta_1 + (\omega - \omega_0)\beta_2)z \qquad (95)$$

This equation tells us that the group delay will be a linear function of frequency, and that higher frequencies will arrive at later times, if β_2 is positive. We refer to the chirp as positive if lower frequencies lead the higher frequencies in time (requiring a positive β_2 in (95)); chirp is negative if the higher frequencies lead in time (negative β_2). Fig. 12.17 shows the broadening effect and illustrates the chirping phenomenon.

✔ **D12.6.** For the fiber channel of Example 12.11, a 20 ps pulse is input instead of the 10 ps pulse in the example. Determine the output pulsewidth.

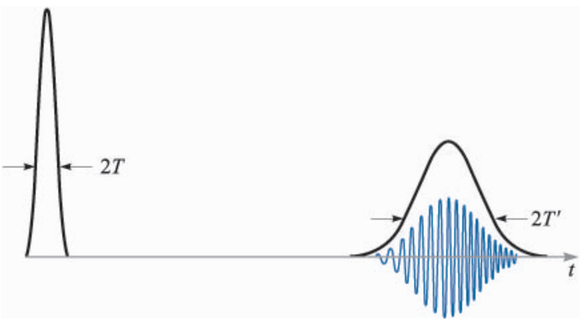


FIGURE 12.17
Gaussian pulse intensities as functions of time (smooth curves) before and after propagation through a dispersive medium, as exemplified by the ω-β diagram of Fig. 12.16*b*. The electric field oscillations are shown under the second trace to demonstrate the chirping effect as the pulse broadens. Note the reduced amplitude of the broadened pulse, which occurs because the pulse energy (the area under the intensity envelope) is constant.

Ans. 25 ps.

As a final point, we note that the pulse bandwidth, $\Delta\omega$, was found to be $1/T$. This is true as long as the Fourier transform of the pulse *envelope* is taken, as was done with (86) to obtain (87). In that case, E_0 was taken to be a constant, and so the only time variation arose from the carrier wave and the Gaussian envelope. Such a pulse, whose frequency spectrum is obtained only from the pulse envelope, is known as *transform-limited*. In general, however, additional frequency bandwidth may be present since E_0 may vary with time for one reason or another (such as phase noise that could be present on the carrier). In these cases, pulse broadening is found from the more general expression

$$\Delta\tau = \Delta\omega\beta_2 z \qquad (96)$$

where $\Delta\omega$ is the net spectral bandwidth arising from all sources. Clearly, transform-limited pulses are preferred in order to minimize broadening, since these will have the smallest spectral width for a given pulse width.

SUGGESTED REFERENCES

1. DuBroff, R. E., S. V. Marshall, and G. G. Skitek: "Electromagnetic Concepts and Applications," 4th ed., Prentice Hall, New Jersey, 1996. Chapter 9 of this text develops the concepts presented here, with additional examples and applications.
2. Iskander, M. F.: "Electromagnetic Fields and Waves," Prentice Hall, New Jersey, 1992. The multiple interface treatment in Chapter 5 of this text is particularly good.
3. Harrington, R. F.: "Time-Harmonic Electromagnetic Fields," McGraw-Hill, New York, 1961. This advanced text provides a good overview of general wave reflection concepts in Chapter 2.
4. Marcuse, D.: "Light Transmission Optics," Van Nostrand Reinhold, New York, 1982. This intermediate-level text provides detailed coverage of optical waveguides and pulse propagation in dispersive media.

PROBLEMS

12.1 A uniform plane wave in air, $E_{x1}^+ = E_{x10}^+ \cos(10^{10}t - \beta z)$ V/m, is normally incident on a copper surface at $z = 0$. What percentage of the incident power density is transmitted into the copper?

12.2 The plane $y = 0$ defines the boundary between two different dielectrics. For $y < 0$, $\epsilon'_{R1} = 1$, $\mu_1 = \mu_0$, and $\epsilon''_{R1} = 0$; and for $y > 0$, $\epsilon'_{R2} = 5$, $\mu_2 = \mu_0$, and $\epsilon''_{R2} = 0$. Let $E_{z1}^+ = 150\cos(\omega t - 8y)$ V/m, and find: (a) ω; (b) $\mathbf{H}_1^+$; (c) $\mathbf{H}_1^-$.

12.3 A uniform plane wave in region 1 is normally incident on the planar boundary separating regions 1 and 2. If $\epsilon''_1 = \epsilon''_2 = 0$, while $\epsilon'_{R1} = \mu^3_{R1}$ and $\epsilon'_{R2} = \mu^3_{R2}$, find the ratio $\epsilon'_{R2}/\epsilon'_{R1}$ if 20% of the energy in the incident wave is reflected at the boundary. There are two possible answers.

12.4 The magnetic field intensity in a region where $\epsilon'' = 0$ is given as $\mathbf{H} = 5\cos\omega t \cos\beta z\,\mathbf{a_y}$ A/m, where $\omega = 5$ Grad/s and $\beta = 30$ rad/m. If the amplitude of the associated electric field intensity is 2 kV/m, find: (a) μ and ϵ' for the medium; (b) $\mathbf{E}$.

12.5 The region $z < 0$ is characterized by $\epsilon'_R = \mu_R = 1$ and $\epsilon''_R = 0$. The total $\mathbf{E}$ field here is given as the sum of two uniform plane waves, $\mathbf{E_s} = 150\,e^{-j10z}\,\mathbf{a_x} + (50\angle 20^\circ)\,e^{j10z}\,\mathbf{a_x}$ V/m. (a) What is the operating frequency? (b) Specify the intrinsic impedance of the region $z > 0$ that would provide the appropriate reflected wave. (c) At what value of z, $-10\,\text{cm} < z < 0$, is the total electric field intensity a maximum amplitude?

12.6 Region 1, $z < 0$, and region 2, $z > 0$, are described by the following parameters: $\epsilon'_1 = 100$ pF/m, $\mu_1 = 25$ μH/m, $\epsilon''_1 = 0$, $\epsilon'_2 = 200$ pF/m, $\mu_2 = 50$ μH/m, and $\epsilon''_2/\epsilon'_2 = 0.5$. If $\mathbf{E}^+_1 = 600\,e^{-\alpha_1 z}\cos(5 \times 10^{10} t - \beta_1 z)$ $\mathbf{a_x}$ V/m, find: (a) α_1; (b) β_1; (c) $\mathbf{E}^+_{s1}$; (d) $\mathbf{E}^-_{s1}$; (e) $\mathbf{E}^+_{s2}$.

12.7 The semi-infinite regions, $z < 0$ and $z > 1$ m, are free space. For $0 < z < 1$ m, $\epsilon'_R = 4$, $\mu_R = 1$, and $\epsilon''_R = 0$. A uniform plane wave with $\omega = 4 \times 10^8$ rad/s is traveling in the $\mathbf{a_z}$ direction toward the interface at $z = 0$. (a) Find the standing wave ratio in each of the three regions. (b) Find the location of the maximum $|\mathbf{E}|$ for $z < 0$ that is nearest to $z = 0$.

12.8 A wave starts at point a, propagates 100 m through a lossy dielectric for which $\alpha = 0.5$ Np/m, reflects at normal incidence at a boundary at which $\Gamma = 0.3 + j0.4$, and then returns to point a. Calculate the ratio of the final power to the incident power after this round trip.

12.9 Region 1, $z < 0$, and region 2, $z > 0$, are both perfect dielectrics ($\mu = \mu_0$, $\epsilon'' = 0$). A uniform plane wave traveling in the $\mathbf{a_z}$ direction has a radian frequency of 3×10^{10} rad/s. Its wavelengths in the two regions are $\lambda_1 = 5$ cm and $\lambda_2 = 3$ cm. What percentage of the energy incident on the boundary is: (a) reflected; (b) transmitted? (c) What is the standing wave ratio in region 1?

12.10 In Fig. 12.1, let region 2 be free space, while $\mu_{R1} = 1$, $\epsilon''_{R1} = 0$, and ϵ'_{R1} is unknown. Find ϵ'_{R1} if: (a) the amplitude of $\mathbf{E}^-_1$ is one-half that of $\mathbf{E}^+_1$; (b) $\mathcal{P}^-_{1,av}$ is one-half of $\mathcal{P}^+_{1,av}$; (c) $|\mathbf{E}_1|_{min}$ is one-half of $|\mathbf{E}_1|_{max}$.

12.11 A 150 MHz uniform plane wave is normally incident from air onto a material whose intrinsic impedance is unknown. Measurements yield a standing wave ratio of 3 and the appearance of an electric field minimum at 0.3 wavelengths in front of the interface. Determine the impedance of the unknown material.

12.12 A 50 MHz uniform plane wave is normally incident from air onto the surface of a calm ocean. For seawater, $\sigma = 4$ S/m, and $\epsilon'_R = 78$. (a) Determine the fractions of the incident power that are reflected and

transmitted. (b) Qualitatively, how will these answers change (if at all) as the frequency is increased?

12.13 A right-circularly polarized plane wave is normally incident from air onto a semi-infinite slab of plexiglas ($\epsilon'_R = 3.45$, $\epsilon''_R = 0$). Calculate the fractions of the incident power that are reflected and transmitted. Also, describe the polarizations of the reflected and transmitted waves.

12.14 A left-circularly polarized plane wave is normally incident onto the surface of a perfect conductor. (a) Construct the superposition of the incident and reflected waves in phasor form. (b) Determine the real instantaneous form of the result of part a. (c) Describe the wave that is formed.

12.15 Consider these regions in which $\epsilon'' = 0$: region 1, $z < 0$, $\mu_1 = 4\ \mu$H/m and $\epsilon'_1 = 10$ pF/m; region 2, $0 < z < 6$ cm, $\mu_2 = 2\ \mu$H/m, $\epsilon'_2 = 25$ pF/m; region 3, $z > 6$ cm, $\mu_3 = \mu_1$ and $\epsilon'_3 = \epsilon'_1$. (a) What is the lowest frequency at which a uniform plane wave incident from region 1 onto the boundary at $z = 0$ will have no reflection? (b) If $f = 50$ MHz, what will the standing wave ratio be in region 1?

12.16 A uniform plane wave in air is normally incident onto a lossless dielectric plate of thickness $\lambda/8$, and of intrinsic impedance $\eta = 260\ \Omega$. Determine the standing wave ratio in front of the plate. Also find the fraction of the incident power that is transmitted to the other side of the plate.

12.17 Repeat Problem 12.16 for the cases in which the frequency is (a) doubled, and (b) quadrupled. Assume that the slab impedance is independent of frequency.

12.18 In Fig. 12.6, let $\eta_1 = \eta_3 = 377\,\Omega$, and $\eta_2 = 0.4\eta_1$. A uniform plane wave is normally incident from the left, as shown. Plot a curve of the standing wave ratio, s in the region to the left: (a) as a function of l if $f = 2.5$ GHz; (b) as a function of frequency if $l = 2$ cm.

12.19 You are given four slabs of lossless dielectric, all with the same intrinsic impedance, η, known to be different from that of free space. The thickness of each slab is $\lambda/4$, where λ is the wavelength as measured in the slab material. The slabs are to be positioned parallel to one another, and the combination lies in the path of a uniform plane wave, normally incident. The slabs are to be arranged such that the air spaces between them are either zero, one-quarter wavelength, or one-half wavelength in thickness. Specify an arrangement of slabs and air spaces such that (a) the wave is totally transmitted through the stack, and (b) the stack presents the highest reflectivity to the incident wave. Several answers may exist.

12.20 The 50 MHz plane wave of Problem 12.12 is incident onto the ocean surface at an angle to the normal of 60°. Determine the fractions of the incident power that are reflected and transmitted for (a) s polarization, and (b) p polarization.

12.21 A right-circularly polarized plane wave in air is incident at Brewster's angle onto a semi-infinite slab of plexiglas ($\epsilon'_R = 3.45$, $\epsilon''_R = 0$). (a) Determine the fractions of the incident power that are reflected and

transmitted. (b) Describe the polarizations of the reflected and transmitted waves.

12.22 A dielectric waveguide is shown in Fig. 12.18 with refractive indices as labeled. Incident light enters the guide at angle ϕ from the front surface normal as shown. Once inside, the light totally reflects at the upper $n_1 - n_2$ interface, where $n_1 > n_2$. All subsequent reflections from the upper and lower boundaries will be total as well, and so the light is confined to the guide. Express, in terms of n_1 and n_2, the maximum value of ϕ such that total confinement will occur, with $n_0 = 1$. The quantity $\sin\phi$ is known as the *numerical aperture* of the guide.

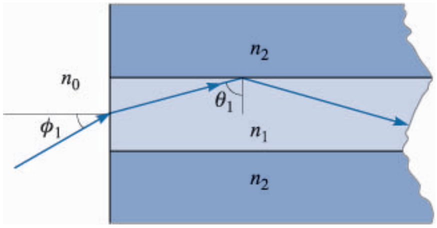


FIGURE 12.18
See Problems 12.22 and 12.23.

12.23 Suppose that ϕ in Fig. 12.18 is Brewster's angle, and that θ_1 is the critical angle. Find n_0 in terms of n_1 and n_2.

12.24 A *Brewster prism* is designed to pass p-polarized light without any reflective loss. The prism of Fig. 12.19 is made of glass ($n = 1.45$), and is in air. Considering the light path shown, determine the apex angle, α.

12.25 In the Brewster prism of Fig. 12.19, determine for s-polarized light the fraction of the incident power that is transmitted through the prism.

12.26 Show how a single block of glass can be used to turn a p-polarized beam of light through 180°, with the light suffering (in principle) zero reflective loss. The light is incident from air, and the returning beam (also in air) may be displaced sideways from the incident beam. Specify all pertinent angles and use $n = 1.45$ for glass. More than one design is possible here.

12.27 Using Eq. (59) in Chapter 11 as a starting point, determine the ratio of the group and phase velocities of an electromagnetic wave in a good conductor. Assume conductivity does not vary with frequency.

12.28 Over a certain frequency range, the refractive index of a certain material varies approximately linearly with frequency:

$$n(\omega) \doteq n_a + n_b(\omega - \omega_a)$$

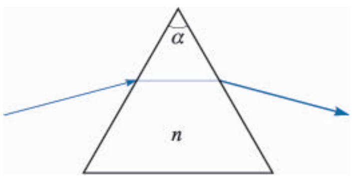


FIGURE 12.19
See Problems 12.24 and 12.25.

where n_a, n_b, and ω_a are constants. Using $\beta = n\omega/c$: (a) determine the group velocity as a function (or perhaps not a function) of frequency; (b) determine the group dispersion parameter, β_2; (c) discuss the implications of these results, if any, on pulse broadening.

12.29 A $T = 5$ ps transform-limited pulse propagates in a dispersive channel for which $\beta_2 = 10\,\text{ps}^2/\text{km}$. Over what distance will the pulse spread to twice its initial width?

12.30 A $T = 20$ ps transform-limited pulse propagates through 10 km of a dispersive channel for which $\beta_2 = 12\,\text{ps}^2/\text{km}$. The pulse then propagates through a second 10 km channel for which $\beta_2 = -12\,\text{ps}^2/\text{km}$. Describe the pulse at the output of the second channel and give a physical explanation for what happened.

CHAPTER 13

TRANSMISSION LINES

Transmission lines are used to transmit electric energy and signals from one point to another, specifically from a source to a load. This may include the connection between a transmitter and an antenna, connections between computers in a network, or between a hydroelectric generating plant and a substation several hundred miles away. Other familiar examples include the interconnects between components of a stereo system, and the connection between a cable service provider and your television set. Examples that are less familiar include the connections between devices on a circuit board that are designed to operate at high frequencies.

What all of the above examples have in common is that the devices to be connected are separated by distances on the order of a wavelength or much larger, whereas in basic circuit analysis methods, connections between elements are of negligible length. The latter condition enabled us, for example, to take for granted that the voltage across a resistor on one side of a circuit was exactly in phase with the voltage source on the other side, or, more generally, that the time measured at the source location is precisely the same time as measured at all other points in the circuit. When distances are sufficiently large between source and receiver, time delay effects become appreciable, leading to the delay-induced phase differences mentioned above. In short, we deal with *wave phenomena* on transmission lines, just as we did with point-to-point energy propagation in free space or in dielectrics.

The basic elements in a circuit, such as resistors, capacitors, inductors, and the connections between them, are considered *lumped* elements if the time delay

in traversing the elements is negligible. On the other hand, if the elements or interconnections are large enough, it may be necessary to consider them as *distributed* elements. This means that their resistive, capacitive, and inductive characteristics must be evaluated on a per-unit-distance basis. Transmission lines have this property in general, and thus become circuit elements in themselves, possessing impedances that contribute to the circuit problem. The basic rule is that one must consider elements as distributed if the propagation delay across the element dimension is on the order of the shortest time interval of interest. In the time-harmonic case, this condition would lead to a measurable phase difference between each end of the device in question.

In this chapter, we investigate wave phenomena in transmission lines, in ways that are very similar to those used in the previous two chapters. Our objectives include (1) to understand how to treat transmission lines as circuit elements possessing complex impedances that are functions of line length and frequency, (2) to understand the properties of different types of lines, (3) to learn methods of combining different transmission lines to accomplish a desired objective, and (4) to understand transient phenomena on lines.

First, however, we need to show that there is a direct analogy between the uniform transmission line and the uniform plane wave. We shall find that the effort devoted to the uniform plane wave in the previous chapters makes it possible to develop analogous results for the uniform transmission line easily and rapidly. The field distributions for the uniform plane wave and for the uniform transmission line are both known as transverse electromagnetic (TEM) waves because **E** and **H** are both perpendicular to the direction of propagation, or both lie in the transverse plane. The great similarity in results is a direct consequence of the fact that we are dealing with TEM waves in each case. In the transmission line, however, it is possible and customary to define a voltage and a current. These quantities are the ones for which we shall write equations, obtain solutions, and find propagation constants, reflection coefficients, and input impedances. We shall also consider power instead of power density.

13.1 THE TRANSMISSION-LINE EQUATIONS

We shall first obtain the differential equations which the voltage or current must satisfy on a uniform transmission line. This may be done by any of several methods. For example, an obvious method would be to solve Maxwell's equations subject to the boundary conditions imposed by the particular transmission line we are considering. We could then define a voltage and a current, thus obtaining our desired equations. It is also possible to solve the general TEM-wave problem once and for all for any two-conductor transmission line having lossless conductors. Instead, we shall construct a circuit model for an incremental length of line, write two circuit equations, and show that the resultant equations are analogous to the fundamental equations from which the wave equation was developed in the previous chapter. By these means we shall begin to tie field theory and circuit theory together.

Our circuit model will contain the inductance, capacitance, shunt conductance, and series resistance associated with an incremental length of line. Let us do our thinking in terms of a coaxial transmission line containing a dielectric of permeability μ (usually μ_0), permittivity ϵ', and conductivity σ.[1] The inner and outer conductors have a high conductivity σ_c. Knowing the operating frequency and the dimensions, we can then determine the values of R, G, L, and C on a per-unit-length basis by using formulas developed in earlier chapters. We shall review these expressions and collect the information on several different types of lines in the following section.

Let us again assume propagation in the $\mathbf{a}_z$ direction. We therefore cut out a section of length Δz containing a resistance $R\Delta z$, an inductance $L\Delta z$, a conductance $G\Delta z$, and a capacitance $C\Delta z$, as shown in Fig. 13.1. Since the section of the line looks the same from either end, we divide the series elements in half to produce a symmetrical network. We could equally well have placed half the conductance and half the capacitance at each end.

Since we are already familiar with the basic characteristics of wave propagation, let us turn immediately to the case of sinusoidal time variation, and use the notation for complex quantities we developed in the last chapter. The voltage V between conductors is in general a function of z and t, as, for example,

$$V(z,t) = V_0 \cos(\omega t - \beta z + \psi)$$

We may use Euler's identity to express this in complex notation,

$$V(z,t) = \text{Re}\left\{V_0 e^{j(\omega t - \beta z + \psi)}\right\} = \text{Re}\left\{V_0 e^{j\psi} e^{-j\beta z} e^{j\omega t}\right\}$$

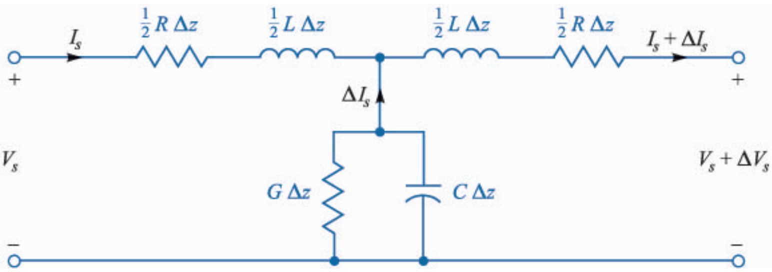


FIGURE 13.1
An incremental length of a uniform transmission line. R, G, L, and C are functions of the transmission-line configuration and materials.

[1] In this basic circuit model, the dielectric loss mechanism is limited to its conductivity, σ. As we considered in Chapter 11, this is a specialization of the more general ϵ'' that characterizes any dielectric loss mechanism (including conductivity), that would be encountered by the fields as they propagate through the line. We retain the notation, ϵ', for the real part of the permittivity.

By dropping Re and suppressing $e^{j\omega t}$, we transform the voltage to a phasor, which we indicate by an s subscript,

$$V_s(z) = V_0 e^{j\psi} e^{-j\beta z}$$

We may now write the voltage equation around the perimeter of the circuit of Fig. 13.1,

$$V_s(z) = \left(\frac{1}{2}R\Delta z + j\frac{1}{2}\omega L\Delta z\right)I_s + \left(\frac{1}{2}R\Delta z + j\frac{1}{2}\omega L\Delta z\right)(I_s + \Delta I_s) + V_s + \Delta V_s$$

or

$$\frac{\Delta V_s}{\Delta z} = -(R + j\omega L)I_s - \left(\frac{1}{2}R + j\frac{1}{2}\omega L\right)\Delta I_s$$

As we let Δz approach zero, ΔI_s also approaches zero, and the second term on the right vanishes. In the limit,

$$\boxed{\frac{dV_s}{dz} = -(R + j\omega L)I_s} \tag{1a}$$

Neglecting second-order effects, we approximate the voltage across the central branch as V_s and obtain a second equation,

$$\frac{\Delta I_s}{\Delta z} \doteq -(G + j\omega C)V_s$$

or

$$\boxed{\frac{dI_s}{dz} = -(G + j\omega C)V_s} \tag{1b}$$

Instead of solving these equations, let us save some time by comparing them with the equations which arise from Maxwell's curl equations for the uniform plane wave in a conducting medium. From

$$\nabla \times \mathbf{E}_s = -j\omega\mu\mathbf{H}_s$$

we set $\mathbf{E}_s = E_{xs}\mathbf{a}_x$ and $\mathbf{H}_s = H_{ys}\mathbf{a}_y$, where E_{xs} and H_{ys} are functions of z only, and obtain a scalar equation that we find to be analogous to Eq. (1a):

$$\frac{dE_{xs}}{dz} = -j\omega\mu H_{ys} \tag{2a}$$

Similarly, from

$$\nabla \times \mathbf{H}_s = (\sigma + j\omega\epsilon')\mathbf{E}_s$$

we have, in analogy to (1b):

$$\frac{dH_{ys}}{dz} = -(\sigma + j\omega\epsilon')E_{xs} \tag{2b}$$

Careful comparison of Eqs. (1*b*) and (2*b*) shows a direct analogy between the following pairs of quantities: I_s and H_{ys}, G and σ, C and ϵ', and V_s and E_{xs}. Replacing the variables in one equation by the corresponding quantities produces the other equation. The analogy is particularly strong in this pair of equations, for the corresponding quantities are measured in almost the same units.

Carrying this same analogy over to Eqs. (1*a*) and (2*a*), we see that it continues to hold and provides one additional analogous pair, L and μ. However, there is also a surprise, for the transmission-line equation is more complicated than the field equation. There is no analog for the conductor resistance per unit length R. Although it would be good salesmanship to say that this shows that field theory is simpler than circuit theory, let us be fair in determining the reason for this omission. Conductor resistance must be determined by obtaining a separate solution to Maxwell's equations within the conductors and forcing the two solutions to satisfy the necessary boundary conditions at the interface. We considered steady current fields in conductors back in Chapter 5, and in Chapter 11, we considered the high-frequency case under the guise of "skin effect"; however, we have looked only briefly at the problem of matching two solutions at the boundary. Thus the term that is omitted in the field equation represents the problem of the fields within the conductors, and the solution of this problem enables us to obtain a value for R in the circuit equation. We maintain the analogy by agreeing to replace $j\omega\mu$ by $R + j\omega L$.[2]

The boundary conditions on V_s and E_{xs} are the same, as are those for I_s and H_{ys}, and thus the solution of our two circuit equations may be obtained from a knowledge of the solution of the two field equations, as obtained in the last chapter. From

$$E_{xs} = E_{x0}e^{-jkz}$$

we obtain the voltage wave

$$V_s = V_0 e^{-\gamma z} \tag{3}$$

where, in a manner consistent with common usage, we have replaced jk for the plane wave with γ, the complex propagation constant for the transmission line. The wave propagates in the $+z$ direction with an amplitude $V_s = V_0$ at $z = 0$ (and $V = V_0$ at $z = 0$, $t = 0$ for $y = 0$). The propagation constant for the uniform plane wave,

$$jk = \sqrt{j\omega\mu(\sigma + j\omega\epsilon')}$$

becomes

$$\boxed{\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)}} \tag{4}$$

[2] When ferrite materials enter the field problem, a complex permeability $\mu = \mu' - j\mu''$ is often used to include the effect of nonohmic losses in that material. Under these special conditions $\omega\mu''$ is analogous to R.

The wavelength is still defined as the distance that provides a phase shift of 2π rad; therefore,

$$\lambda = \frac{2\pi}{\beta} \tag{5}$$

Also, the phase velocity has been defined as

$$v_p = \frac{\omega}{\beta} \tag{6}$$

and this expression is valid both for the uniform plane wave and transmission lines. For a lossless line ($R = G = 0$) we see that

$$\gamma = j\beta = j\omega\sqrt{LC}$$

Hence

$$v_p = \frac{1}{\sqrt{LC}} \tag{7}$$

From the expression for the magnetic field intensity

$$H_{ys} = \frac{E_{x0}}{\eta} e^{-jkz}$$

we see that the positively traveling current wave

$$I_s = \frac{V_0}{Z_0} e^{-\gamma z} \tag{8}$$

is related to the positively traveling voltage wave by a *characteristic impedance* Z_0 that is analogous to η. Since, in a conducting medium

$$\eta = \sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon'}}$$

we have

$$\boxed{Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}} \tag{9}$$

When a uniform plane wave in medium 1 is incident on the interface with medium 2, the fraction of the incident wave that is reflected is called the reflection coefficient, Γ, which for normal incidence is

$$\Gamma = \frac{E_{x0}^-}{E_{x0}^+} = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1}$$

Thus the fraction of the incident voltage wave that is reflected by a line with a different characteristic impedance, say Z_{02}, is

$$\Gamma = |\Gamma|e^{j\phi} = \frac{V_0^-}{V_0^+} = \frac{Z_{02} - Z_{01}}{Z_{02} + Z_{01}} \tag{10}$$

Knowing the reflection coefficient, we may find the standing-wave ratio,

$$s = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{11}$$

Finally, when $\eta = \eta_3$ for $z > 0$, and $\eta = \eta_2$ for $z < 0$, the ratio of E_{xs} to H_{ys} at $z = -l$ is

$$\eta_{in} = \eta_2 \frac{\eta_3 \cos \beta_2 l + j\eta_2 \sin \beta_2 l}{\eta_2 \cos \beta_2 l + j\eta_3 \sin \beta_2 l}$$

and therefore the input impedance

$$Z_{in} = Z_{02} \frac{Z_{03} \cos \beta_2 l + jZ_{02} \sin \beta_2 l}{Z_{02} \cos \beta_2 l + jZ_{03} \sin \beta_2 l} \tag{12}$$

is the ratio of V_s to I_s at $z = -l$ when $Z_0 = Z_{03}$ for $z > 0$ and is Z_{02} for $z < 0$. We often terminate a transmission line at $z = 0$ with a load impedance Z_L which may represent an antenna, the input circuit of a television receiver, or an amplifier on a telephone line. The input impedance at $z = -l$ is then written simply as

$$Z_{in} = Z_0 \frac{Z_L \cos \beta l + jZ_0 \sin \beta l}{Z_0 \cos \beta l + jZ_L \sin \beta l} \tag{13}$$

Let us illustrate the use of several of these transmission line formulas with a basic example.

Example 13.1

A lossless transmission line is 80 cm long and operates at a frequency of 600 MHz. The line parameters are $L = 0.25\,\mu\text{H/m}$ and $C = 100$ pF/m. Find the characteristic impedance, the phase constant, the velocity on the line, and the input impedance for $Z_L = 100\ \Omega$.

Solution. Since the line is lossless, both R and G are zero. The characteristic impedance is

$$Z_0 = \sqrt{\frac{L}{C}} = \sqrt{\frac{0.25 \times 10^{-6}}{100 \times 10^{-12}}} = 50\ \Omega$$

Since $\gamma = \alpha + j\beta = \sqrt{(R + j\omega L)(G + j\omega C)} = j\omega\sqrt{LC}$, we see that

$$\beta = \omega\sqrt{LC} = 2\pi(600 \times 10^6)\sqrt{(0.25 \times 10^{-6})(100 \times 10^{-12})} = 18.85\ \text{rad/m}$$

Also,

$$v_p = \frac{\omega}{\beta} = \frac{2\pi(600 \times 10^6)}{18.85} = 2 \times 10^8 \text{ m/s}$$

We now have all the necesssary information to find Z_{in} from (13):

$$Z_{in} = Z_0 \frac{Z_L \cos\beta l + jZ_0 \sin\beta l}{Z_0 \cos\beta l + jZ_L \sin\beta l} = 50 \frac{100\cos(18.85 \times 0.8) + j50\sin(18.85 \times 0.8)}{50\cos(18.85 \times 0.8) + j100\sin(18.85 \times 0.8)}$$
$$= 60.3\angle 35.5^\circ = 49.1 + j35.0\,\Omega$$

✔ **D13.1.** At an operating radian frequency of 500 Mrad/s, typical circuit values for a certain transmission line are: $R = 0.2\,\Omega/\text{m}$, $L = 0.25\,\mu\text{H/m}$, $G = 10\,\mu\text{S/m}$, and $C = 100\,\text{pF/m}$. Find: (a) α; (b) β; (c) λ; (d) v_p; (e) Z_0.

Ans. 2.25 mNp/m; 2.50 rad/m; 2.51 m; 2×10^8 m/sec; $50.0 - j0.0350\,\Omega$.

13.2 TRANSMISSION-LINE PARAMETERS

Let us use this section to collect previous results and develop new ones where necessary, so that values for R, G, L, and C are available for the simpler types of transmission lines.

Coaxial (High Frequencies)

We begin by seeing how many of the necessary expressions we already have for a coaxial cable in which the dielectric has an inner radius a and outer radius b (Fig. 13.2). The capacitance per unit length, obtained as Eq. (46) of Sec. 5.10, is

$$C = \frac{2\pi\epsilon'}{\ln(b/a)} \tag{14}$$

The value of permittivity used should be appropriate for the range of operating frequencies considered.

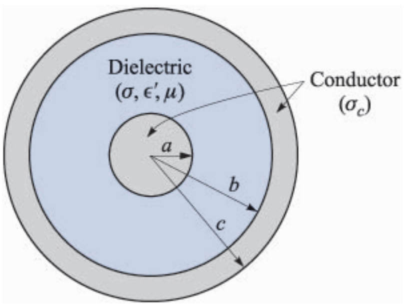


FIGURE 13.2
The geometry of the coaxial transmission line. A homogeneous dielectric is assumed.

The conductance per unit length may be determined easily from the capacitance expression above by use of the current analogy described in Sec. 6.3. Thus,

$$G = \frac{2\pi\sigma}{\ln(b/a)} \tag{15}$$

where σ is the conductivity of the dielectric between the conductors at the operating frequency.

The inductance per unit length was computed for the coaxial cable as Eq. (50) in Sec. 9.10,

$$L_{ext} = \frac{\mu}{2\pi}\ln(b/a) \tag{16}$$

where μ is the permeability of the dielectric between conductors, usually μ_0. This is an *external* inductance, for its calculation does not take into account any flux within either conductor. Equation (16) is usually an excellent approximation to the total inductance of a high-frequency transmission line, however, for the skin depth is so small at typical operating frequencies that there is negligible flux within either conductor and negligible internal inductance. Note that $L_{ext}C = \mu\epsilon' = 1/v_p^2$, and we are therefore able to evaluate the external inductance for any transmission line for which we know the capacitance and insulator characteristics.

The last of the four parameters that we need is the resistance R per unit length. If the frequency is very high and the skin depth δ is very small, then we obtain an appropriate expression for R by distributing the total current uniformly throughout a depth δ. For a circular conductor of radius a and conductivity σ_c, we let Eq. (54) of Sec. 11.5 apply to a unit length, obtaining

$$R_{inner} = \frac{1}{2\pi a\delta\sigma_c}$$

There is also a resistance for the outer conductor, which has an inner radius b. We assume the same conductivity σ_c and the same value of skin depth δ, leading to

$$R_{outer} = \frac{1}{2\pi b\delta\sigma_c}$$

Since the line current flows through these two resistances in series, the total resistance is the sum:

$$R = \frac{1}{2\pi\delta\sigma_c}\left(\frac{1}{a} + \frac{1}{b}\right) \tag{17}$$

It is convenient to include the common expression for the characteristic impedance of a coax here with the parameter formulas. Thus

$$Z_0 = \sqrt{\frac{L_{ext}}{C}} = \frac{1}{2\pi}\sqrt{\frac{\mu}{\epsilon'}}\ln\frac{b}{a} \tag{18}$$

If necessary, a more accurate value may be obtained from (9).

Coaxial (Low Frequencies)

Now let us spend a few paragraphs obtaining the parameter values at very low frequencies where there is no appreciable skin effect and the current is assumed to be distributed uniformly throughout the cross section.

We first note that the current distribution in the conductor does not affect either the capacitance or conductance per unit length. Hence

$$C = \frac{2\pi\epsilon'}{\ln(b/a)} \tag{14}$$

and

$$G = \frac{2\pi\sigma}{\ln(b/a)} \tag{15}$$

The resistance per unit length may be calculated by dc methods, $R = l/(\sigma_c S)$, where $l = 1\text{m}$ and σ_c is the conductivity of the outer and inner conductors. The area of the center conductor is πa^2 and that of the outer is $\pi(c^2 - b^2)$. Adding the two resistance values, we have

$$R = \frac{1}{\sigma_c \pi}\left(\frac{1}{a^2} + \frac{1}{c^2 - b^2}\right) \tag{19}$$

Only one of the four parameter values remains to be found, the inductance per unit length. The external inductance that we calculated at high frequencies is the greatest part of the total inductance. However, smaller terms must be added to it, representing the internal inductances of the inner and outer conductors.

At very low frequencies where the current distribution is uniform, the internal inductance of the center conductor is the subject of Prob. 43 in Chap. 9; the relationship is also given as Eq. (62) in Sec. 9.10:

$$L_{a,int} = \frac{\mu}{8\pi} \tag{20}$$

The determination of the internal inductance of the outer shell is a more difficult problem, and most of the work is requested in Prob. 7 at the end of this chapter. There, we find that the energy stored per unit length in an outer cylindrical shell of inner radius b and outer radius c with uniform current distribution is

$$W_H = \frac{\mu I^2}{16\pi(c^2 - b^2)}\left(b^2 - 3c^2 + \frac{4c^2}{c^2 - b^2}\ln\frac{c}{b}\right)$$

Thus the internal inductance of the outer conductor at very low frequencies is

$$L_{bc,int} = \frac{\mu}{8\pi(c^2 - b^2)}\left(b^2 - 3c^2 + \frac{4c^2}{c^2 - b^2}\ln\frac{c}{b}\right) \tag{21}$$

At low frequencies the total inductance is obtained by combining (16), (20), and (21):

$$L = \frac{\mu}{2\pi}\left[\ln\frac{b}{a} + \frac{1}{4} + \frac{1}{4(c^2 - b^2)}\left(b^2 - 3c^2 + \frac{4c^2}{c^2 - b^2}\ln\frac{c}{b}\right)\right] \tag{22}$$

Coaxial (Intermediate Frequencies)

There still remains the frequency interval where the skin depth is neither very much larger than nor very much smaller than the radius. In this case, the current distribution is governed by Bessel functions, and both the resistance and internal inductance are complicated expressions. Values are tabulated in the handbooks, and it is necessary to use them for very small conductor sizes at high frequencies and for larger conductor sizes used in power transmission at low frequencies.[3]

Two-Wire (High Frequencies)

For the two-wire transmission line of Fig. 13.3 with conductors of radius a and conductivity σ_c with center-to-center separation d in a medium of permeability μ, permittivity ϵ', and conductivity σ_c, the capacitance was found in Sec. 5.11 to be

$$C = \frac{\pi\epsilon'}{\cosh^{-1}(d/2a)} \tag{23}$$

or

$$C \doteq \frac{\pi\epsilon'}{\ln(d/a)} \qquad (a << d)$$

The external inductance may be found from $L_{ext}C = \mu\epsilon'$. It is

$$L_{ext} = \frac{\mu}{\pi}\cosh^{-1}(d/2a) \tag{24}$$

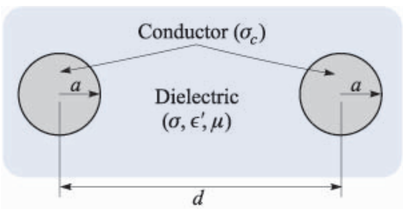


FIGURE 13.3
The geometry of the two-wire transmission line.

[3] The current distribution, internal inductance, and internal resistance of round wires is discussed (with numerical examples) in Weeks, pp. 35–44. See the Suggested References at the end of this chapter.

or

$$L_{ext} \doteq \frac{\mu}{\pi} \ln(d/a) \qquad (a << d)$$

The conductance per unit length may be written immediately from an inspection of the capacitance expression,

$$G = \frac{\pi\sigma}{\cosh^{-1}(d/2a)} \tag{25}$$

The resistance per unit length is twice that of the center conductor of the coax,

$$R = \frac{1}{\pi a \delta \sigma_c} \tag{26}$$

Finally, using the capacitance and the external inductance expressions, we obtain a value for the characteristic impedance for the lossless case ($R = G = 0$),

$$Z_0 = \sqrt{\frac{L_{ext}}{C}} = \frac{1}{\pi}\sqrt{\frac{\mu}{\epsilon'}} \cosh^{-1}(d/2a) \tag{27}$$

Two-Wire (Low Frequencies)

At low frequencies where a uniform current distribution may be assumed, we again must modify the L and R expressions. We therefore have the same relationships for C and G:

$$C = \frac{\pi\epsilon'}{\cosh^{-1}(d/2a)} \tag{23}$$

$$G = \frac{\pi\sigma}{\cosh^{-1}(d/2a)} \tag{25}$$

but the inductance per unit length must be increased by twice the internal inductance of a straight round wire,

$$L = \frac{\mu}{\pi}\left[\frac{1}{4} + \cosh^{-1}(d/2a)\right] \tag{28}$$

and the resistance becomes twice the dc resistance of a wire of radius a, conductivity σ_c, and unit length:

$$R = \frac{2}{\pi a^2 \sigma_c} \tag{29}$$

Planar (High Frequencies)

If we have the parallel-plane or planar transmission line of Fig. 13.4, with two conducting planes of conductivity σ_c, thickness t, separation d, and a dielectric with parameters ϵ', μ, and σ, then we may easily determine the circuit parameters

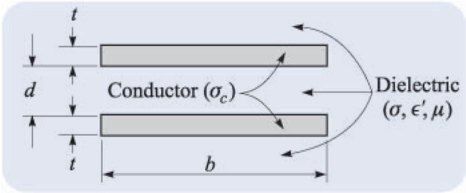


FIGURE 13.4
The geometry of the planar transmission line.

per unit length for a width b. It is necessary to assume either that $b >> d$ or that we are considering a width b of a much wider guiding system. We have

$$C = \frac{\epsilon' b}{d} \tag{30}$$

$$G = \frac{\sigma b}{d} \tag{31}$$

$$L_{ext} = \frac{\mu d}{b} \tag{32}$$

and

$$R = \frac{2}{\sigma_c \delta b} \tag{33}$$

Here we have assumed a well-developed skin effect such that $d << t$, the thickness of either plane.

Finally, for the lossless line,

$$Z_0 = \sqrt{\frac{L_{ext}}{C}} = \sqrt{\frac{\mu}{\epsilon'}}\frac{d}{b} \tag{34}$$

Low-frequency parameter values are unimportant since the planar transmission line has little use there.

✔ **D13.2.** The dimensions of a coaxial transmission line are $a = 4$ mm, $b = 17.5$ mm, and $c = 20$ mm. The conductivity of the inner and outer conductors is 2×10^7 S/m, and the dielectric properties are $\mu_R = 1$, $\epsilon'_R = 3$, and $\sigma/\omega\epsilon' = 0.025$. Assume that the loss tangent is constant with frequency. Determine: (a) L, C, R, G, and Z_0 at 150 MHz; (b) L and R at 60 Hz.

Ans. 0.295 μH/m, 113.1 pF/m, 0.266 Ω/m, 2.66 mS/m, 51.1 Ω; 0.355 μH/m, 1.164 mΩ/m.

✔ **D13.3.** The conductors of a two-wire transmission line each have a radius of 0.8 mm and a conductivity of 3×10^7 S/m. They are separated a center-to-center distance of 0.8 cm in a medium for which $\epsilon'_R = 2.5$, $\mu_R = 1$, and $\sigma = 4 \times 10^{-9}$ S/m. If the line operates at 60 Hz, find: (a) δ; (b) C; (c) G; (d) L; (e) R.

Ans. 1.186 cm; 30.3 pF/m; 5.48 nS/m; 1.017 μH/m; 0.0332 Ω/m.

✔ **D13.4.** Parameters for the planar transmission line shown in Fig. 13.4 are $b = 6$ mm, $d = 0.25$ mm, $t = 25$ mm, $\sigma_c = 5.5 \times 10^7$ S/m, $\epsilon' = 25$ pF/m, $\mu = \mu_0$, and $\sigma/\omega\epsilon' = 0.03$. If the operating frequency is 750 MHz, calculate: (a) α; (b) β; (c) Z_0.

Ans. 0.470 Np/m; 26.4 rad/m; $9.34\angle 0.699^\circ$ Ω.

13.3 SOME TRANSMISSION-LINE EXAMPLES

In this section we shall apply many of the results that we have obtained in the previous two sections to several typical transmission-line problems. We shall simplify our work by restricting our attention to the lossless line.

Let us begin by assuming a two-wire 300-Ω line ($Z_0 = 300\,\Omega$), such as the lead-in wire from the antenna to a television or FM receiver. The circuit is shown in Fig. 13.5. The line is 2 m long and the dielectric constant is such that the velocity on the line is 2.5×10^8 m/s. We shall terminate the line with a receiver having an input resistance of 300 Ω and represent the antenna by its Thevenin equivalent $Z_{Th} = 300\,\Omega$ in series with $V_{s,Th} = 60$ V at 100 MHz. This antenna voltage is larger by a factor of about 10^5 than it would be in a practical case, but it also provides simpler values to work with; in order to think practical thoughts, divide currents or voltages by 10^5, divide powers by 10^{10}, and leave impedances alone.

Since the load impedance is equal to the characteristic impedance, the line is matched; the reflection coefficient is zero, and the standing wave ratio is unity. For the given velocity and frequency, the wavelength on the line is $v/f = 2.5$ m, and the phase constant is $2\pi/\lambda = 0.8\pi$ rad/m; the attenuation constant is zero. The electrical length of the line is $\beta l = (0.8\pi)2$, or 1.6π rad. This length may also be expressed as 288°, or 0.8 wavelength.

The input impedance offered to the voltage source is 300 Ω, and since the internal impedance of the source is 300 Ω, the voltage at the input to the line is half of 60 V, or 30 V. The source is matched to the line and delivers the maximum available power to the line. Since there is no reflection and no attenuation, the voltage at the load is 30 V, but it is delayed in phase by 1.6π rad. Thus

$$V_{in} = 30\cos(2\pi 10^8 t) \quad \text{V}$$

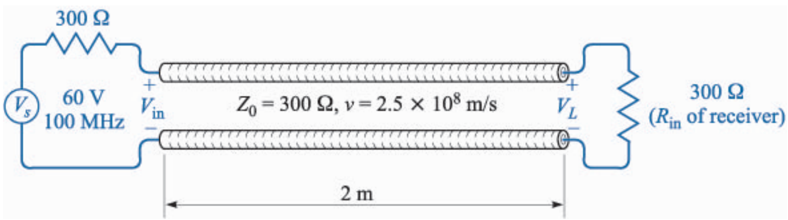


FIGURE 13.5
A transmission line that is matched at each end produces no reflections and thus delivers maximum power to the load.

whereas

$$V_L = 30\cos(2\pi 10^8 t - 1.6\pi) \quad \text{V}$$

The input current is

$$I_{in} = \frac{V_{in}}{300} = 0.1\cos(2\pi 10^8 t) \quad \text{A}$$

while the load current is

$$I_L = 0.1\cos(2\pi 10^8 t - 1.6\pi) \quad \text{A}$$

The average power delivered to the input of the line by the source must all be delivered to the load by the line,

$$P_{in} = P_L = \frac{1}{2} \times 30 \times 0.1 = 1.5 \quad \text{W}$$

Now let us connect a second receiver, also having an input resistance of 300 Ω, across the line in parallel with the first receiver. The load impedance is now 150 Ω, the reflection coefficient is

$$\Gamma = \frac{150 - 300}{150 + 300} = -\frac{1}{3}$$

and the standing wave ratio on the line is

$$s = \frac{1 + \frac{1}{3}}{1 - \frac{1}{3}} = 2$$

The input impedance is no longer 300 Ω, but is now

$$Z_{in} = Z_0 \frac{Z_L \cos\beta l + jZ_0 \sin\beta l}{Z_0 \cos\beta l + jZ_L \sin\beta l} = 300\,\frac{150\cos 288^\circ + j300\sin 288^\circ}{300\cos 288^\circ + j150\sin 288^\circ}$$
$$= 510\angle{-23.8^\circ} = 466 - j206 \quad \Omega$$

which is a capacitive impedance. Physically, this means that this length of line stores more energy in its electric field than in its magnetic field. The input current phasor is thus

$$I_{s,in} = \frac{60}{300 + 466 - j206} = 0.0756\angle 15.0^\circ \quad \text{A}$$

and the power supplied to the line by the source is

$$P_{in} = \frac{1}{2} \times (0.0756)^2 \times 466 = 1.333 \quad \text{W}$$

Since there are no losses in the line, 1.333 W must also be delivered to the load. Note that this is less than the 1.50 W which we were able to deliver to a matched load; moreover, this power must divide equally between two receivers, and thus each receiver now receives only 0.667 W. Since the input impedance of each receiver is 300 Ω, the voltage across the receiver is easily found as

$$0.667 = \frac{1}{2}\frac{|V_{s,L}|^2}{300}$$

$$|V_{s,L}| = 20 \quad \text{V}$$

in comparison with the 30 V obtained across the single load.

Before we leave this example, let us ask ourselves several questions about the voltages on the transmission line. Where is the voltage a maximum and a minimum, and what are these values? Does the phase of the load voltage still differ from the input voltage by 288°? Presumably, if we can answer these questions for the voltage, we could do the same for the current.

We answered questions of this nature for the uniform plane wave in the last chapter, and our analogy should therefore provide us with the corresponding information for the transmission line. In Sec. 12.2, Eq. (21) serves to locate the voltage maxima at

$$z_{max} = -\frac{1}{2\beta}(\phi + 2m\pi) \quad (m = 0, 1, 2, \ldots)$$

where $\Gamma = |\Gamma|e^{j\phi}$. Thus, with $\beta = 0.8\pi$ and $\phi = \pi$, we find

$$z_{max} = -0.625 \text{ and } -1.875 \quad \text{m}$$

while the minima are $\lambda/4$ distant from the maxima,

$$z_{min} = 0 \text{ and } -1.25 \quad \text{m}$$

and we find that the load voltage (at $z = 0$) is a voltage minimum. This, of course, verifies the general conclusion we reached in the last chapter: a voltage minimum occurs at the load if $Z_L < Z_0$, and a voltage maximum occurs if $Z_L > Z_0$, where both impedances are pure resistances.

The minimum voltage on the line is thus the load voltage, 20 V; the maximum voltage must be 40 V, since the standing wave ratio is 2. The voltage at the input end of the line is

$$V_{s,in} = I_{s,in}Z_{in} = (0.0756\angle 15.0°)(510\angle -23.8°) = 38.5\angle -8.8°$$

The input voltage is almost as large as the maximum voltage anywhere on the line because the line is about three-quarters wavelength long, a length which would place the voltage maximum at the input when $Z_L < Z_0$.

The final question we posed for ourselves deals with the relative phase of the input and load voltages. Although we have found each of these voltages, we do not know the phase angle of the load voltage. From Sec. 12.2, Eq. (18), the voltage at any point on the line is

$$V_s = \left(e^{-j\beta z} + \Gamma e^{j\beta z}\right)V_0^+ \tag{35}$$

We may use this expression to determine the voltage at any point on the line in terms of the voltage at any other point. Since we know the voltage at the input to the line, we let $z = -l$,

$$V_{s,in} = \left(e^{j\beta l} + \Gamma e^{-j\beta l}\right) V_0^+ \tag{36}$$

and solve for V_0^+,

$$V_0^+ = \frac{V_{s,in}}{e^{j\beta l} + \Gamma e^{-j\beta l}} = \frac{38.5\angle{-8.8^\circ}}{e^{j1.6\pi} - \frac{1}{3}e^{-j1.6\pi}} = 30.0\angle 72.0^\circ \quad \text{V}$$

We may now let $z = 0$ in (35) to find the load voltage,

$$V_{s,L} = (1 + \Gamma)V_0^+ = 20\angle 72^\circ = 20\angle{-288^\circ}$$

The amplitude agrees with our previous value. The presence of the reflected wave causes $V_{s,in}$ and $V_{s,L}$ to differ in phase by about -279° instead of -288°.

Example 13.2

In order to provide a slightly more complicated example, let us now place a purely capacitive impedance of $-j300\,\Omega$ in parallel with the two 300-Ω receivers. We are to find the input impedance and the power delivered to each receiver.

Solution. The load impedance is now 150 Ω in parallel with $-j300\,\Omega$, or

$$Z_L = \frac{150(-j300)}{150 - j300} = \frac{-j300}{1 - j2} = 120 - j60 \quad \Omega$$

We first calculate the reflection coefficient and the standing wave ratio:

$$\Gamma = \frac{120 - j60 - 300}{120 - j60 + 300} = \frac{-180 - j60}{420 - j60} = 0.447\angle{-153.4^\circ}$$

$$s = \frac{1 + 0.447}{1 - 0.447} = 2.62$$

Thus, the standing wave ratio is higher and the mismatch is therefore worse. Let us next calculate the input impedance. The electrical length of the line is still 288°, so that

$$Z_{in} = 300\,\frac{(120 - j60)\cos 288^\circ + j300 \sin 288^\circ}{300 \cos 288^\circ + j(120 - j60)\sin 288^\circ} = 755 - j138.5 \quad \Omega$$

This leads to a source current of

$$I_{s,in} = \frac{V_{Th}}{Z_{Th} + Z_{in}} = \frac{60}{300 + 755 - j138.5} = 0.0564\angle 7.47^\circ \quad \text{A}$$

Therefore, the average power delivered to the input of the line is $P_{in} = \frac{1}{2}(0.0564)^2(755) = 1.200$ W. Since the line is lossless, it follows that $P_L = 1.200$ W, and each receiver gets only 0.6 W.

Example 13.3

As a final example let us terminate our line with a purely capacitive impedance, $Z_L = -j300\,\Omega$. We seek the reflection coefficient, the standing-wave ratio, and the power delivered to the load.

Solution. Obviously, we cannot deliver any average power to the load since it is a pure reactance. As a consequence, the reflection coefficient is

$$\Gamma = \frac{-j300 - 300}{-j300 + 300} = -j1 = 1\angle{-90^\circ}$$

and the reflected wave is equal in amplitude to the incident wave. Hence it should not surprise us to see that the standing wave ratio is

$$s = \frac{1 + |-j1|}{1 - |-j1|} = \infty$$

and the input impedance is a pure reactance,

$$Z_{in} = 300\,\frac{-j300\cos 288^\circ + j300\sin 288^\circ}{300\cos 288^\circ + j(-j300)\sin 288^\circ} = j589$$

Thus, no average power can be delivered to the input impedance by the source, and therefore no average power can be delivered to the load.

Although we could continue to find numerous other facts and figures for these examples, much of the work may be done more easily for problems of this type by using graphical techniques. We shall encounter these in the following section.

D13.5. A 50-W lossless line has a length of 0.4λ. The operating frequency is 300 MHz. A load $Z_L = 40 + j30\ \Omega$ is connected at $z = 0$, and the Thevenin-equivalent source at $z = -l$ is $12\angle 0^\circ$ V in series with $Z_{Th} = 50 + j0\ \Omega$. Find: (a) Γ; (b) s; (c) Z_{in}.

Ans. $0.333\angle 90^\circ$; 2.00; $25.5 + j5.90\ \Omega$

D13.6. For the transmission line of Prob. D13.5, also find: (a) the phasor voltage at $z = -l$; (b) the phasor voltage at $z = 0$; (c) the average power delivered to Z_L.

Ans $4.14\angle 8.58^\circ$ V; $6.32\angle{-125.6^\circ}$ V; 0.320 W.

13.4 GRAPHICAL METHODS

Transmission line problems often involve manipulations with complex numbers, making the time and effort required for a solution several times greater than that needed for a similar sequence of operations on real numbers. One means of reducing the labor without seriously affecting the accuracy is by using transmission-line charts. Probably the most widely used one is the Smith chart.[4]

Basically, this diagram shows curves of constant resistance and constant reactance; these may represent either an input impedance or a load impedance. The latter, of course, is the input impedance of a zero-length line. An indication of location along the line is also provided, usually in terms of the fraction of a

[4] P.H. Smith, "Transmission Line Calculator," *Electronics*, vol. 12, pp. 29–31, January, 1939.

wavelength from a voltage maximum or minimum. Although they are not specifically shown on the chart, the standing-wave ratio and the magnitude and angle of the reflection coefficient are very quickly determined. As a matter of fact, the diagram is constructed within a circle of unit radius, using polar coordinates, with radius variable $|\Gamma|$ and counterclockwise angle variable ϕ, where $\Gamma = |\Gamma|e^{j\phi}$. Figure 13.6 shows this circle. Since $|\Gamma| < 1$, all our information must lie on or within the unit circle. Peculiarly enough, the reflection coefficient itself will not be plotted on the final chart, for these additional contours would make the chart very difficult to read.

The basic relationship upon which the chart is constructed is

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{37}$$

The impedances which we plot on the chart will be *normalized* with respect to the characteristic impedance. Let us identify the normalized load impedance as z_L,

$$z_L = r + jx = \frac{Z_L}{Z_0} = \frac{R_L + jX_L}{Z_0}$$

and thus

$$\Gamma = \frac{z_L - 1}{z_L + 1}$$

or

$$z_L = \frac{1 + \Gamma}{1 - \Gamma} \tag{38}$$

In polar form, we have used $|\Gamma|$ and ϕ as the magnitude and angle of Γ; let us now select Γ_r and Γ_i as the real and imaginary parts of Γ,

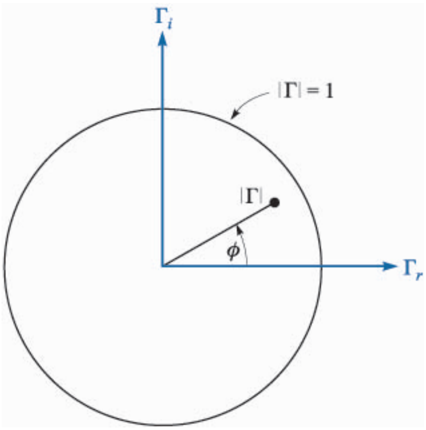


FIGURE 13.6
The polar coordinates of the Smith chart are the magnitude and phase angle of the reflection coefficient; the cartesian coordinates are the real and imaginary parts of the reflection coefficient. The entire chart lies within the unit circle $|\Gamma| = 1$.

$$\Gamma = \Gamma_r + j\Gamma_i \tag{39}$$

Thus

$$r + jx = \frac{1 + \Gamma_r + j\Gamma_i}{1 - \Gamma_r - j\Gamma_i} \tag{40}$$

The real and imaginary parts of this equation are

$$r = \frac{1 - \Gamma_r^2 - \Gamma_i^2}{(1 - \Gamma_r)^2 + \Gamma_i^2} \tag{41}$$

$$x = \frac{2\Gamma_i}{(1 - \Gamma_r)^2 + \Gamma_i^2} \tag{42}$$

After several lines of elementary algebra, we may write (41) and (42) in forms which readily display the nature of the curves on Γ_r, Γ_i axes,

$$\left(\Gamma_r - \frac{r}{1+r}\right)^2 + \Gamma_i^2 = \left(\frac{1}{1+r}\right)^2 \tag{43}$$

$$(\Gamma_r - 1)^2 + \left(\Gamma_i - \frac{1}{x}\right)^2 = \left(\frac{1}{x}\right)^2 \tag{44}$$

The first equation describes a family of circles, where each circle is associated with a specific value of resistance r. For example, if $r = 0$ the radius of this zero-resistance circle is seen to be unity, and it is centered at $\Gamma_r = 0$, $\Gamma_i = 0$, the origin. This checks, for a pure reactance termination leads to a reflection coefficient of unity magnitude. On the other hand, if $r = \infty$, then $z_L = \infty$ and we have $\Gamma = 1 + j0$. The circle described by (43) is centered at $\Gamma_r = 1$, $\Gamma_i = 0$ and has zero radius. It is therefore the point $\Gamma = 1 + j0$, as we decided it should be. As another example, the circle for $r = 1$ is centered at $\Gamma_r = 0.5$, $\Gamma_i = 0$ and has a radius of 0.5. This circle is shown on Fig. 13.7, along with circles for $r = 0.5$ and $r = 2$. All circles are centered on the Γ_r axis and pass through the point $\Gamma = 1 + j0$.

Equation (44) also represents a family of circles, but each of these circles is defined by a particular value of x, rather than r. If $x = \infty$, then $z_L = \infty$, and $\Gamma = 1 + j0$ again. The circle described by (44) is centered at $\Gamma = 1 + j0$ and has zero radius; it is therefore the point $\Gamma = 1 + j0$. If $x = +1$, then the circle is

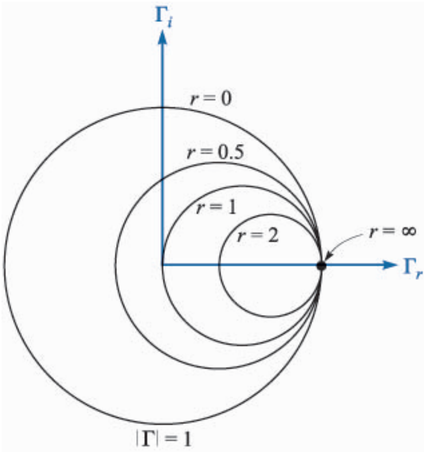


FIGURE 13.7
Constant-r circles are shown on the Γ_r, Γ_i plane. The radius of any circle is $1/(1 + r)$.

centered at $\Gamma = 1 + j1$ and has unit radius. Only one-quarter of this circle lies within the boundary curve $|\Gamma| = 1$, as shown in Fig. 13.8. A similar quarter-circle appears below the Γ_r axis for $x = -1$. The portions of other circles for $x = 0.5$, -0.5, 2, and -2 are also shown. The "circle" representing $x = 0$ is the Γ_r axis; this is also labeled on Fig. 13.8.

The two families of circles both appear on the Smith chart, as shown in Fig. 13.9. It is now evident that if we are given Z_L, we may divide by Z_0 to obtain z_L, locate the appropriate r and x circles (interpolating as necessary), and determine Γ by the intersection of the two circles. Since the chart does not have concentric circles showing the values of $|\Gamma|$, it is necessary to measure the radial distance from the origin to the intersection with dividers or compass and use an auxiliary scale to find $|\Gamma|$. The graduated line segment below the chart in Fig. 13.9 serves this purpose. The angle of Γ is ϕ, and it is the counter-clockwise angle from the Γ_r axis. Again, radial lines showing the angle would clutter up the chart badly, so

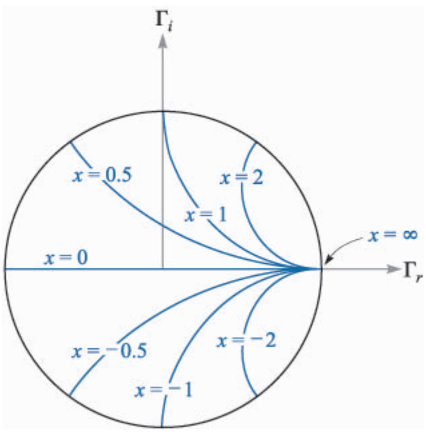


FIGURE 13.8
The portions of the circles of constant x lying within $|\Gamma| = 1$ are shown on the Γ_r, Γ_i axes. The radius of a given circle is $1/|x|$.

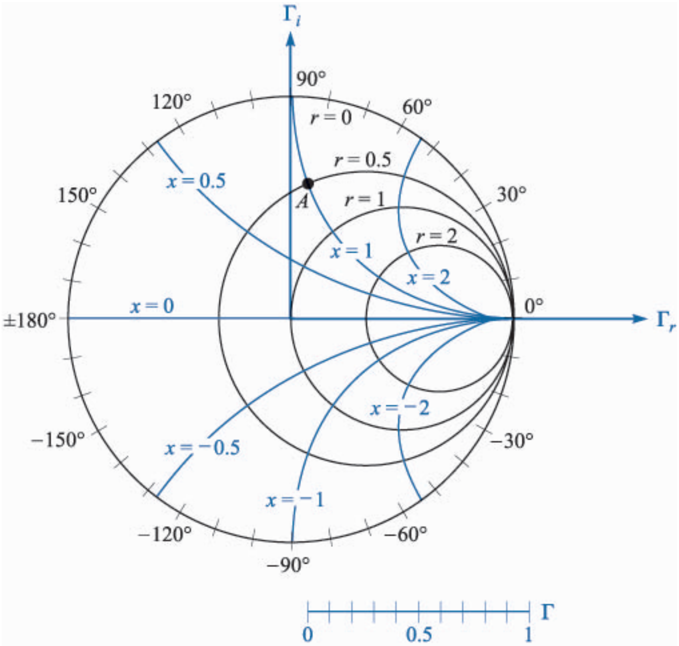


FIGURE 13.9
The Smith chart contains the constant-r circles and constant-x circles, an auxiliary radial scale to determine $|\Gamma|$, and an angular scale on the circumference for measuring ϕ.

the angle is indicated on the circumference of the circle. A straight line from the origin through the intersection may be extended to the perimeter of the chart. As an example, if $Z_L = 25 + j50\ \Omega$ on a 50-Ω line, $z_L = 0.5 + j1$, and point A on Fig. 13.9 shows the intersection of the $r = 0.5$ and $x = 1$ circles. The reflection coefficient is approximately 0.62 at an angle ϕ of 83°.

The Smith chart is completed by adding a second scale on the circumference by which distance along the line may be computed. This scale is in wavelength units, but the values placed on it are not obvious. To obtain them, we first divide the voltage at any point along the line,

$$V_s = V_0^+\left(e^{-j\beta z} + \Gamma e^{j\beta z}\right)$$

by the current

$$I_s = \frac{V_0^+}{Z_0}\left(e^{-j\beta z} - \Gamma e^{j\beta z}\right)$$

obtaining the normalized input impedance

$$z_{in} = \frac{V_s}{Z_0 I_s} = \frac{e^{-j\beta z} + \Gamma e^{j\beta z}}{e^{-j\beta z} - \Gamma e^{j\beta z}}$$

Replacing z by $-l$ and dividing numerator and denominator by $e^{j\beta l}$, we have the general equation relating normalized input impedance, reflection coefficient, and line length,

$$z_{in} = \frac{1 + \Gamma e^{-j2\beta l}}{1 - \Gamma e^{-j2\beta l}} = \frac{1 + |\Gamma| e^{j(\phi - 2\beta l)}}{1 - |\Gamma| e^{j(\phi - 2\beta l)}} \tag{45}$$

Note that when $l = 0$, we are located at the load, and $z_{in} = (1 + \Gamma)/(l - \Gamma) = z_L$, as shown by (38).

Equation (45) shows that the input impedance at any point $z = -l$ can be obtained by replacing Γ, the reflection coefficient of the load, by $\Gamma e^{-j2\beta l}$. That is, we decrease the angle of Γ by $2\beta l$ radians as we move from the load to the line input. Only the angle of Γ is changed; the magnitude remains constant.

Thus, as we proceed from the load z_L to the input impedance z_{in}, we move *toward* the generator a distance l on the transmission line, but we move through a *clockwise* angle of $2\beta l$ on the Smith chart. Since the magnitude of Γ stays constant, the movement toward the source is made along a constant-radius circle. One lap around the chart is accomplished whenever βl changes by π rad, or when l changes by one-half wavelength. This agrees with our earlier discovery that the input impedance of a half-wavelength lossless line is equal to the load impedance.

The Smith chart is thus completed by the addition of a scale showing a change of 0.5λ for one circumnavigation of the unit circle. For convenience, two scales are usually given, one showing an increase in distance for clockwise movement and the other an increase for counterclockwise travel. These two scales are shown in Fig. 13.10. Note that the one marked "wavelengths toward generator" (wtg) shows increasing values of l/λ for clockwise travel, as described above. The zero point of the wtg scale is rather arbitrarily located to the left. This corresponds to input impedances having phase angles of 0° and $R_L < Z_0$. We have also seen that voltage minima are always located here.

Example 13.4

The use of the transmission line chart is best shown by example. Let us again consider a load impedance, $Z_L = 25 + j50\ \Omega$, terminating a 50-Ω line. The line length is 60 cm and the operating frequency is such that the wavelength on the line is 2 m. We desire the input impedance.

Solution. We have $z_L = 0.5 + j1$, which is marked as A on Fig. 13.11, and we read $\Gamma = 0.62\angle 82^\circ$. By drawing a straight line from the origin through A to the circumference, we note a reading of 0.135 on the wtg scale. We have $l/\lambda = 0.6/2 = 0.3$, and it is therefore 0.3λ from the load to the input. We therefore find z_{in} on the $|\Gamma| = 0.62$ circle opposite a wtg reading of $0.135 + 0.300 = 0.435$. This construction is shown in Fig. 13.11, and the point locating the input impedance is marked B. The normalized input impedance is read as $0.28 - j0.40$, and thus $Z_{in} = 14 - j20$. A more accurate analytical calculation gives $Z_{in} = 13.7 - j20.2$.

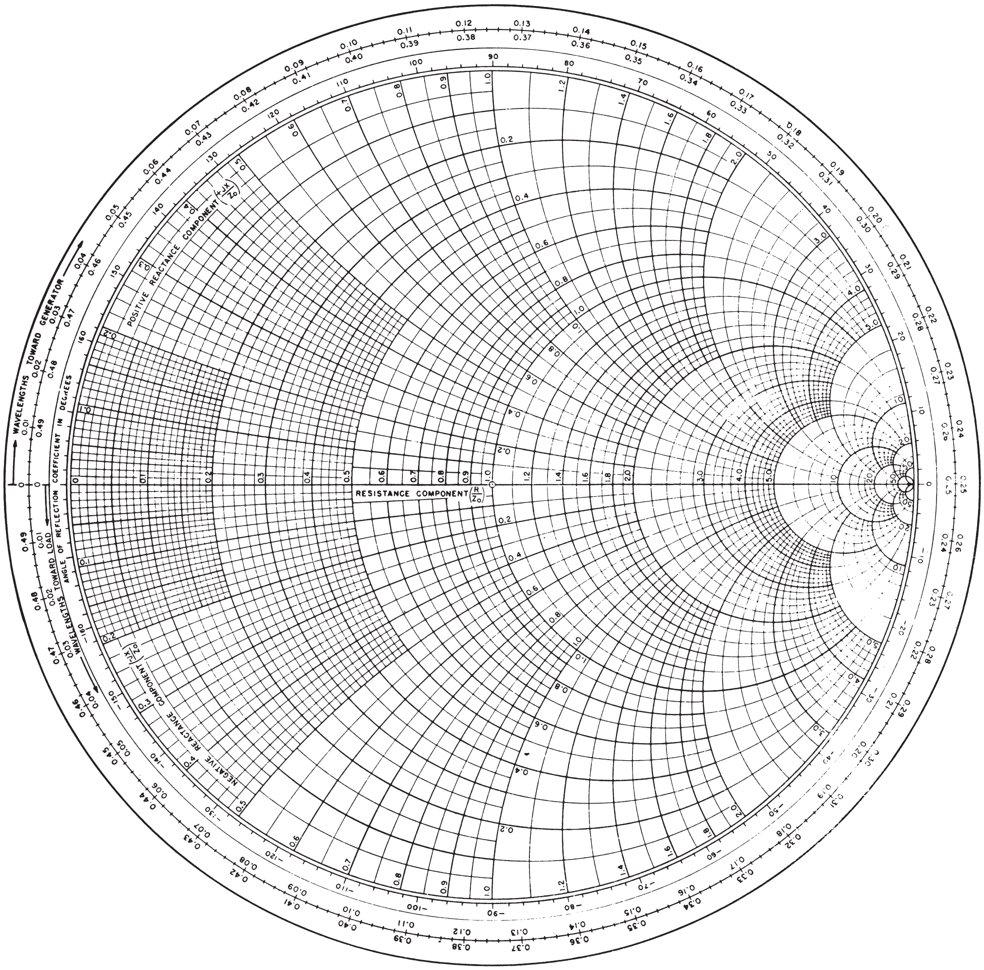


FIGURE 13.10
A photographic reduction of one version of a useful Smith chart (*courtesy of the Emeloid Company, Hillside, N.J.*). For accurate work, larger charts are available wherever fine technical books are sold.

Information concerning the location of the voltage maxima and minima is also readily obtained on the Smith chart. We already know that a maximum or minimum must occur at the load when Z_L is a pure resistance; if $R_L > Z_0$ there is a maximum at the load, and if $R_L < Z_0$ there is a minimum. We may extend this result now by noting that we could cut off the load end of a transmission line at a point where the input impedance is a pure resistance and replace that section with a resistance R_{in}; there would be no changes on the generator portion of the line. It follows, then, that the location of voltage maxima and minima must be at those points where Z_{in} is a pure resistance. Purely resistive input impedances

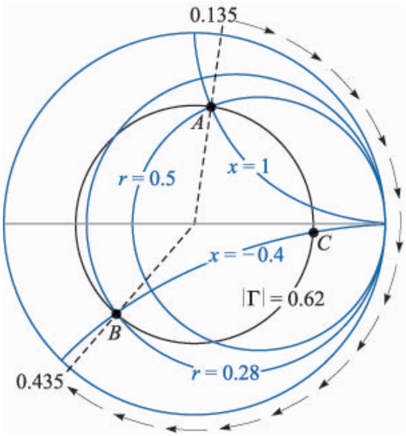


FIGURE 13.11
The normalized input impedance produced by a normalized load impedance $z_L = 0.5 + j1$ on a line 0.3λ long is $z_{in} = 0.28 - j0.40$.

must occur on the $x = 0$ line (the Γ_r axis) of the Smith chart. Voltage maxima or current minima occur when $r > 1$, or at wtg $= 0.25$, and voltage minima or current maxima occur when $r < 1$, or at wtg $= 0$. In the example above, then, the maximum at wtg $= 0.250$ must occur $0.250 - 0.135 = 0.115$ wavelengths toward the generator from the load. This is a distance of 0.115×200, or 23 cm from the load.

We should also note that since the standing wave ratio produced by a resistive load R_L is either R_L/R_0 or R_0/R_L, whichever is greater than unity, the value of s may be read directly as the value of r at the intersection of the $|\Gamma|$ circle and the r axis, $r > 1$. In our example this intersection is marked point C, and $r = 4.2$; thus, $s = 4.2$.

Transmission line charts may also be used for normalized admittances, although there are several slight differences in such use. We let $y_L = Y_L/Y_0 = g + jb$ and use the r circles as g circles and the x circles as b circles. The two differences are: first, the line segment where $g > 1$ and $b = 0$ corresponds to a voltage minimum; and second, 180° must be added to the angle of Γ as read from the perimeter of the chart. We shall use the Smith chart in this way in the following section.

Special charts are also available for non-normalized lines, particularly 50-Ω charts and 20-mS charts.

✔ **D13.7.** A load $Z_L = 80 - j100\ \Omega$ is located at $z = 0$ on a lossless 50-Ω line. The operating frequency is 200 MHz and the wavelength on the line is 2 m. (a) If the line is 0.8 m in length, use the Smith chart to find the input impedance. (b) What is s? (c) What is the distance from the load to the nearest voltage maximum? (d) What is the distance from the input to the nearest point at which the remainder of the line could be replaced by a pure resistance?

Ans. $79 + j99\ \Omega$: 4.50; 0.0397 m; 0.760 m

13.5 SEVERAL PRACTICAL PROBLEMS

In this section we shall direct our attention to two examples of practical transmission line problems. The first is the determination of load impedance from experimental data, and the second is the design of a single-stub matching network.

Let us assume that we have made experimental measurements on a 50-Ω air line which show that there is a standing wave ratio of 2.5. This has been determined by moving a sliding carriage back and forth along the line to determine maximum and minimum voltage readings. A scale provided on the track along which the carriage moves indicates that a *minimum* occurs at a scale reading of 47.0 cm, as shown in Fig. 13.12. The zero point of the scale is arbitrary and does not correspond to the location of the load. The location of the minimum is usually specified instead of the maximum because it can be determined more accurately than that of the maximum; think of the sharper minima on a rectified sine wave. The frequency of operation is 400 MHz, so the wavelength is 75 cm. In order to pinpoint the location of the load, we remove it and replace it with a short circuit; the position of the minimum is then determined as 26.0 cm.

We know that the short circuit must be located an integral number of half-wavelengths from the minimum; let us arbitrarily locate it one half-wavelength away at $26.0 - 37.5 = -11.5$ cm on the scale. Since the short circuit has replaced the load, the load is also located at -11.5 cm. Our data thus show that the minimum is $47.0 - (-11.5) = 58.5$ cm from the load, or subtracting one-half wavelength, a minimum is 21.0 cm from the load. The voltage *maximum* is thus $21.0 - (37.5/2) = 2.25$ cm from the load, or $2.25/75 = 0.030$ wavelength from the load.

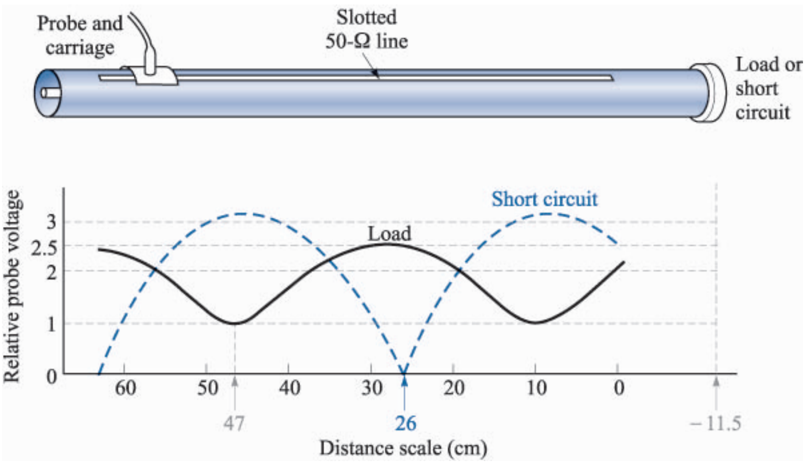


FIGURE 13.12
A sketch of a coaxial slotted line. The distance scale is on the slotted line. With the load in place, $s = 2.5$, and the minimum occurs at a scale reading of 47 cm; for a short circuit the minimum is located at a scale reading of 26 cm. The wavelength is 75 cm.

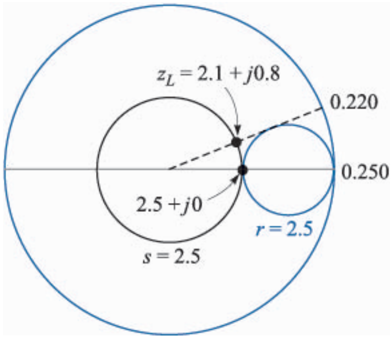


FIGURE 13.13
If $z_{in} = 2.5 + j0$ on a line 0.03 wavelength long, then $z_L = 2.1 + j0.8$.

With this information, we can now turn to the Smith chart. At a voltage maximum the input impedance is a pure resistance equal to sR_0; on a normalized basis, $z_{in} = 2.5$. We therefore enter the chart at $z_{in} = 2.5$ and read 0.250 on the wtg scale. Subtracting 0.030 wavelength to reach the load, we find that the intersection of the $s = 2.5$ (or $|\Gamma| = 0.429$) circle and the radial line to 0.220 wavelength is at $z_L = 2.1 + j0.8$. The construction is sketched on the Smith chart of Fig. 13.13. Thus $Z_L = 105 + j40\,\Omega$, a value which assumes its location at a scale reading of -11.5 cm, or an integral number of half-wavelengths from that position. Of course, we may select the "location" of our load at will by placing the short circuit at that point which we wish to consider as the load location. Since load locations are not well defined, it is important to specify the point (or plane) at which the load impedance is determined.

As a final example, let us try to match this load to the 50-Ω line by placing a short-circuited stub of length d_1 a distance d from the load (see Fig. 13.14). The stub line has the same characteristic impedance as the main line. The lengths d and d_1 are to be determined.

The input impedance to the stub is a pure reactance; when combined in parallel with the input impedance of the length d containing the load, the resul-

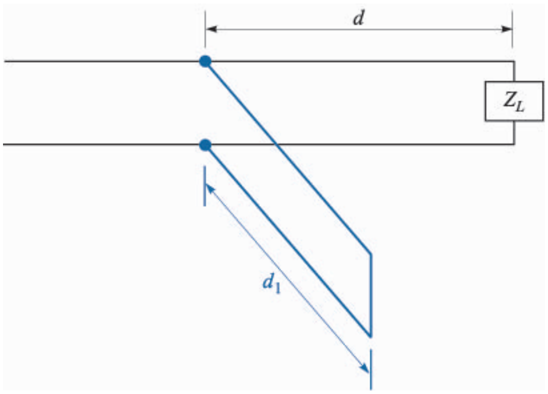


FIGURE 13.14
A short-circuited stub of length d_l, located a distance d from a load Z_L is used to provide a matched load to the left of the stub.

tant input impedance must be $1 + j0$. Since it is much easier to combine admittances in parallel than impedances, let us rephrase our goal in admittance language: the input admittance of the length d containing the load must be $1 + jb_{in}$ for the addition of the input admittance of the stub jb_{stub} to produce a total admittance of $1 + j0$. Hence the stub admittance is $-jb_{in}$. We shall therefore use the Smith chart as an admittance chart instead of an impedance chart.

The impedance of the load is $2.1 + j0.8$, and its location is at -11.5 cm. The admittance of the load is therefore $1/(2.1 + j0.8)$, and this value may be determined by adding one-quarter wavelength on the Smith chart, since Z_{in} for a quarter-wavelength line is R_0^2/Z_L, or $z_{in} = 1/z_L$, or $y_{in} = z_L$. Entering the chart (Fig. 13.15) at $z_L = 2.1 + j0.8$, we read 0.220 on the wtg scale; we add (or subtract) 0.250 and find the admittance $0.41 - j0.16$ corresponding to this impedance. This point is still located on the $s = 2.5$ circle. Now, at what point or points on this circle is the real part of the admittance equal to unity? There are two answers, $1 + j0.95$ at wtg = 0.16, and $1 - j0.95$ at wtg = 0.34, as shown in Fig. 13.15. Let us select the former value since this leads to the shorter stub. Hence $y_{stub} = -j0.95$, and the stub location corresponds to wtg = 0.16. Since the load admittance was found at wtg = 0.470, then we must move $(0.5 - 0.47) + 0.16 = 0.19$ wavelength to get to the stub location.

Finally, we may use the chart to determine the necessary length of the short-circuited stub. The input conductance is zero for any length of short-circuited stub, so we are restricted to the perimeter of the chart. At the short circuit, $y = \infty$ and wtg = 0.250. We find that $b_{in} = -0.95$ is achieved at wtg = 0.379, as shown in Fig. 13.15. The stub is therefore $0.379 - 0.250 = 0.129$ wavelength, or 9.67 cm long.

✔ **D13.8.** Standing wave measurements on a lossless 75-Ω line show maxima of 18 V and minima of 5 V. One minimum is located at a scale reading of 30 cm. With the load replaced by a short circuit, two adjacent minima are found at scale readings of 17 and 37 cm. Find: (a) s; (b) λ; (c) f; (d) Γ_L; (e) Z_L.

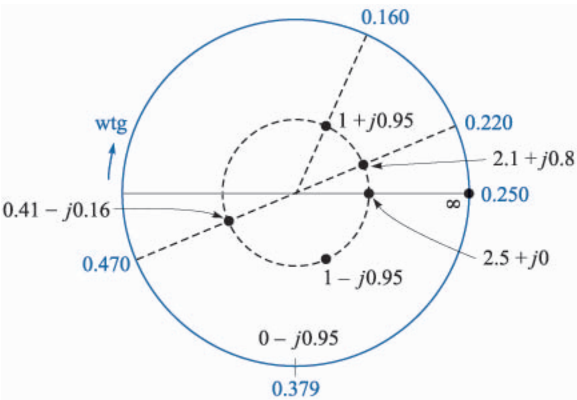


FIGURE 13.15
A normalized load $z_L = 2.1 + j0.8$ is matched by placing a 0.129-wavelength short-circuited stub 0.19 wavelength from the load.

Ans. 3.60; 0.400 m; 750 MHz; $0.704\angle{-33.0}$; $77.9 + j104.7\,\Omega$

✔ **D13.9.** A normalized load, $z_L = 2 - j1$, is located at $z = 0$ on a lossless 50-Ω line. Let the wavelength be 100 cm. (a) A short-circuited stub is to be located at $z = -d$. What is the shortest suitable value for d? (b) What is the shortest possible length of the stub? Find s: (c) on the main line for $z < -d$; (d) on the main line for $-d < z < 0$; (e) on the stub.

Ans. 12.5 cm; 12.5 cm; 1.00; 2.62; ∞.

13.6 TRANSIENTS ON TRANSMISSION LINES

Throughout this chapter, we have considered the operation of transmission lines under steady state conditions, in which voltage and current were sinusoidal and at a single frequency. In this section we move away from the simple time-harmonic case and consider transmission line responses to voltage step functions and pulses, grouped under the general heading of *transients*. Line operation in transient mode is important to study, as it allows us to understand how lines can be used to store and release energy (in pulse-forming applications, for example). Pulse propagation is important in general since digital signals, composed of sequences of pulses, are widely used.

We will confine our discussion to the propagation of transients in lines that are lossless and have no dispersion, so that the basic behavior and analysis methods may be learned. We must remember, however, that transient signals are necessarily composed of numerous frequencies, as Fourier analysis will show. Consequently, the question of dispersion in the line arises, since, as we have found, line propagation constants and reflection coefficients at complex loads will be frequency-dependent. So in general, pulses are likely to broaden with propagation distance, and pulse shapes may change when reflecting from a complex load. These issues will not be considered in detail here, but are readily addressed when the precise frequency dependences of β and Γ are known. In particular, $\beta(\omega)$ can be found by evaluating the imaginary part of γ, as given in Eq. (4), which would in general include the frequency dependences of R, C, G, and L arising from various mechanisms. For example, the skin effect (which affects both the conductor resistance and the internal inductance) will result in frequency-dependent R and L. Once $\beta(\omega)$ is known, pulse broadening can be evaluated using the methods presented in Chapter 12.

We begin our basic discussion of transients by considering a lossless transmission line of length, l, terminated by a matched load, $R_L = Z_0$, as shown in Fig. 13.16a. At the front end of the line is a battery of voltage, V_0, which is connected to the line by closing a switch. At time $t = 0$, the switch is closed, and the line voltage at $z = 0$ becomes equal to the battery voltage. This voltage, however, does not appear across the load until adequate time has elapsed for the propagation delay. Specifically, at $t = 0$, a voltage wave is initiated in the line

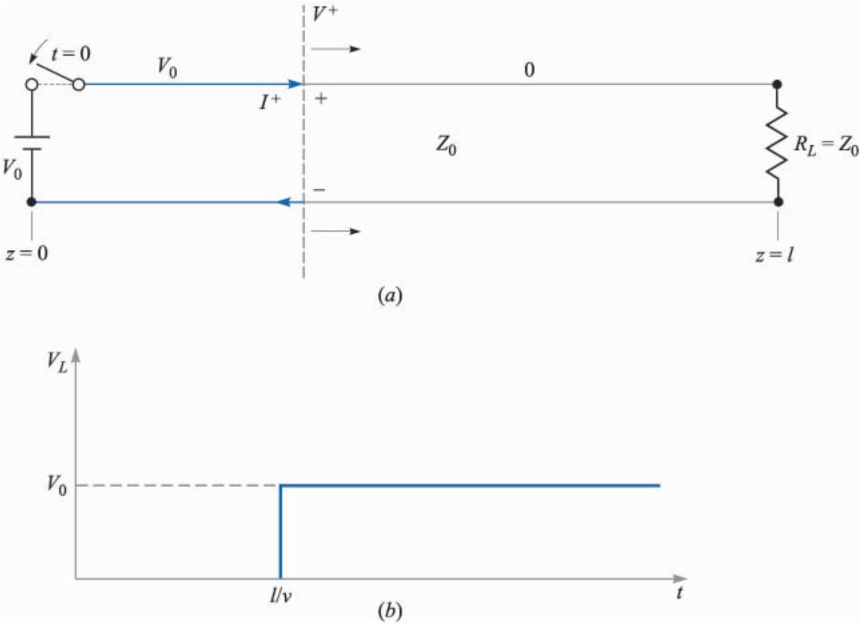


FIGURE 13.16
(a) Closing the switch at time $t = 0$ initiates voltage and current waves, V^+ and I^+. The leading edge of both waves is indicated by the dashed line, which propagates in the lossless line toward the load at velocity v. In this case, $V^+ = V_0$; the line voltage is V^+ everywhere to the left of the leading edge, where current is $I^+ = V^+/Z_0$. To the right of the leading edge, voltage and current are both zero. Clockwise current, indicated here, is treated as positive, and will occur when V^+ is positive. (b) Voltage across the load resistor as a function of time, showing the one-way transit time delay (l/v).

at the battery end, which then propagates toward the load. The leading edge of the wave, labeled V^+ in the figure, is of value $V^+ = V_0$. It can be thought of as a propagating step function, since at all points to the left of V^+, the line voltage is V_0; at all points to the right (not yet reached by the leading edge), the line voltage is zero. The wave propagates at velocity v, which in general is the group velocity in the line.[5] The wave reaches the load at time $t = l/v$, and then does not reflect, since the load is matched. The transient phase is thus over, and the load voltage is equal to the battery voltage. A plot of load voltage as a function of time is shown in Fig. 13.16b, indicating the propagation delay of $t = l/v$.

[5] Since we have a step function (composed of many frequencies) as opposed to a sinusoid at a single frequency, the wave will propagate at the group velocity. In a lossless line with no dispersion as considered in this section, $\beta = \omega\sqrt{LC}$, where L and C are constant with frequency. In this case we would find that the group and phase velocities are equal; (i.e, $d\omega/d\beta = \omega/\beta = v = 1/\sqrt{LC}$). We will thus write the velocity as v, knowing it to be both v_p and v_g.

Associated with the voltage wave, V^+, is a current wave whose leading edge is of value I^+. This wave is a propagating step function as well, whose value at all points to the left of V^+ is $I^+ = V^+/Z_0$; at all points to the right, current is zero. A plot of current through the load as a function of time will thus be identical in form to the voltage plot of Fig. 13.16*b*, except that the load current at $t = l/v$ will be $I_L = V^+/Z_0 = V_0/R_L$.

We next consider a more general case, in which the load of Fig. 13.16*a* is again a resistor, but is *not matched* to the line ($R_L \neq Z_0$). Reflections will occur at the load, thus complicating the problem. At $t = 0$, the switch is closed as before and a voltage wave, $V_1^+ = V_0$, propagates to the right. Upon reaching the load, however, the wave will now reflect, producing a back-propagating wave, V_1^-. The relation between V_1^- and V_1^+ is through the reflection coefficient at the load:

$$\frac{V_1^-}{V_1^+} = \Gamma_L = \frac{R_L - Z_0}{R_L + Z_0} \tag{46}$$

As V_1^- propagates back toward the battery, it leaves behind its leading edge a total voltage of $V_1^+ + V_1^-$. Voltage V_1^+ exists everywhere ahead of the V_1^- wave until it reaches the battery, whereupon the entire line now is charged to voltage $V_1^+ + V_1^-$. At the battery, the V_1^- wave reflects to produce a new forward wave, V_2^+. The ratio of V_2^+ and V_1^- is found through the reflection coefficient at the battery:

$$\frac{V_2^+}{V_1^-} = \Gamma_g = \frac{Z_g - Z_0}{Z_g + Z_0} = \frac{0 - Z_0}{0 + Z_0} = -1 \tag{47}$$

where the impedance at the generator end, Z_g, is that of the battery, or zero.

V_2^+ (equal to $-V_1^-$) now propagates to the load, where it reflects to produce backward wave $V_2^- = \Gamma_L V_2^+$. This wave then returns to the battery, where it reflects with $\Gamma_g = -1$, and the process repeats. Note that with each round trip the wave voltage is reduced in magnitude since $|\Gamma_L| < 1$. Because of this the propagating wave voltages will eventually approach zero, and steady state is reached.

The voltage across the load resistor can be found at any given time by summing the voltage waves that have reached the load and have reflected from it up to that time. After many round trips, the load voltage will be in general:

$$\begin{aligned} V_L &= V_1^+ + V_1^- + V_2^+ + V_2^- + V_3^+ + V_3^- + \dots \\ &= V_1^+\left(1 + \Gamma_L + \Gamma_g\Gamma_L + \Gamma_g\Gamma_L^2 + \Gamma_g^2\Gamma_L^2 + \Gamma_g^2\Gamma_L^3 + \dots\right) \end{aligned}$$

Performing a simple factoring operation, the above becomes

$$V_L = V_1^+(1 + \Gamma_L)\left(1 + \Gamma_g\Gamma_L + \Gamma_g^2\Gamma_L^2 + \dots\right) \tag{48}$$

Allowing time to approach infinity, the second term in parenthesis in (48) becomes the power series expansion for the expression $1/(1 - \Gamma_g \Gamma_L)$. Thus, in steady state we obtain,

$$V_L = V_1^+ \left(\frac{1 + \Gamma_L}{1 - \Gamma_g \Gamma_L} \right) \tag{49}$$

In our present example, $V_1^+ = V_0$ and $\Gamma_g = -1$. Substituting these into (49), we find the expected result in steady state: $V_L = V_0$.

A more general situation would involve a non zero impedance at the battery location, as shown in Fig. 13.17. In this case, a resistor of value R_g is positioned in series with the battery. When the switch is closed, the battery voltage appears across the series combination of R_g and the line characteristic impedance, Z_0. The value of the initial voltage wave, V_1^+, is thus found through simple voltage division, or

$$V_1^+ = \frac{V_0 Z_0}{R_g + Z_0} \tag{50}$$

With this initial value, the sequence of reflections and the development of the voltage across the load occurs in the same manner as determined by (48), with the steady state value determined by (49). The value of the reflection coefficient at the generator end, determined by (47), is $\Gamma_g = (R_g - Z_0)/(R_g + Z_0)$.

A useful way of keeping track of the voltage at any point in the line is through a *voltage reflection diagram*. Such a diagram for the line of Fig. 13.17 is shown in Fig. 13.18*a*. It is a two-dimensional plot in which position on the line, z, is shown on the horizontal axis. Time is plotted on the vertical axis, and is conveniently expressed as it relates to position and velocity through $t = z/v$. A vertical line, located at $z = l$, is drawn which, together with the ordinate, define the z axis boundaries of the transmission line. With the switch located at the

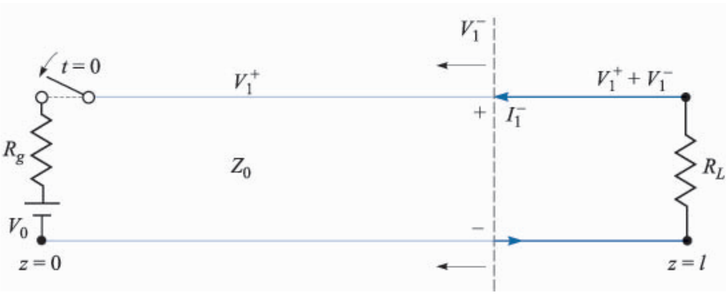
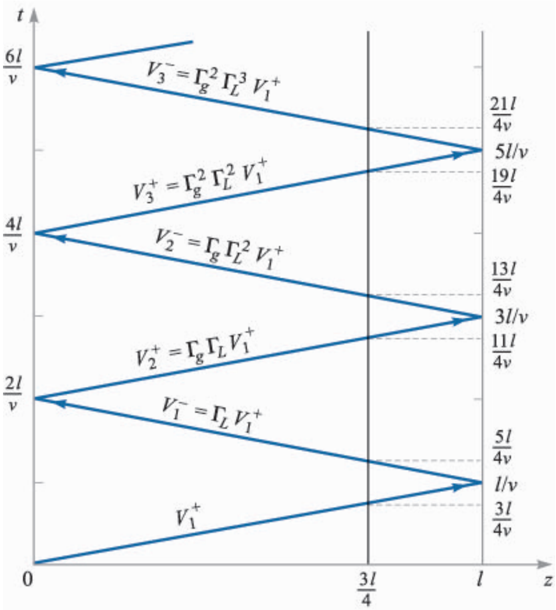
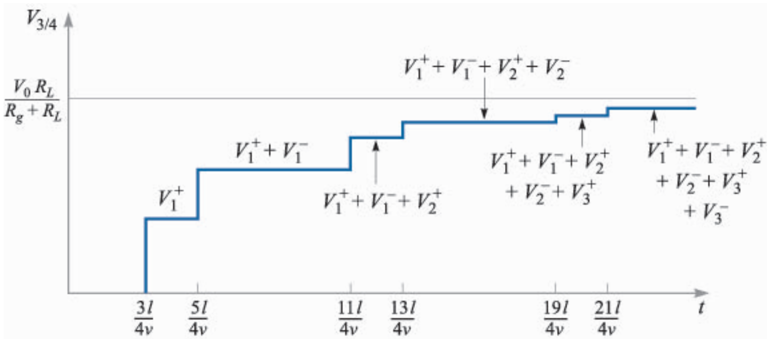


FIGURE 13.17
With a series resistance at the battery location, voltage division occurs when the switch is closed, such that $V_0 = V_{Rg} + V_1^+$. Shown is the first reflected wave, which leaves voltage $V_1^+ + V_1^-$ behind its leading edge. Associated with the wave is current I_1^-, which is $-V_1^-/Z_0$. Counter-clockwise current is treated as negative, and will occur when V_1^- is positive.



(a)



(b)

FIGURE 13.18
(a) Voltage reflection diagram for the line of Fig. 13.17. A reference line, drawn at $z = 3l/4$, is used to evaluate the voltage at that position as a function of time. (b) The line voltage at $z = 3l/4$ as determined from the reflection diagram of (a). Note that the voltage approaches the expected $V_0 R_L/(R_g + R_L)$ as time approaches infinity.

battery position, the initial voltage wave, V_1^+, starts at the origin, or lower left corner of the diagram ($z = t = 0$). The location of the leading edge of V_1^+ as a function of time is shown as the diagonal line that joins the origin to the point along the right-hand vertical line that corresponds to time $t = l/v$ (the one-way

transit time). From there (the load location) the position of the leading edge of the reflected wave, V_1^-, is shown as a "reflected" line which joins the $t = l/v$ point on the right boundary to the $t = 2l/v$ point on the ordinate. From there (at the battery location) the wave reflects again, forming V_2^+, shown as a line parallel to that for V_1^+. Subsequent reflected waves are shown, and their values are labeled.

The voltage as a function of time at a given position in the line can now be determined by adding the voltages in the waves as they intersect a vertical line, drawn at the desired location. This addition is performed starting at the bottom of the diagram ($t = 0$) and progressing upward (in time). Whenever a voltage wave crosses the vertical line, its value is added to the total at that time. For example, the voltage at a location three-fourths the distance from the battery to the load is plotted in Fig. 13.18*b*. To obtain this plot, the line $z = (3/4)l$ is drawn on the diagram. Whenever a wave crosses this line, the voltage in the wave is added to the voltage that has accumulated at $z = (3/4)l$ over all earlier times. This general procedure enables one to easily determine the voltage at any specific time and location. In doing so, the terms in (48) that have occurred up to the chosen time are being added, but with information on the time at which each term appears.

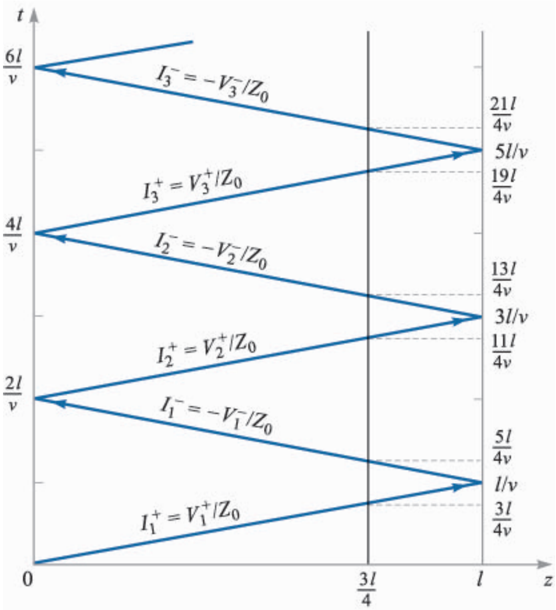
Line current can be found in a similar way through a *current reflection diagram*. It is easiest to construct the current diagram directly from the voltage diagram by determining a value for current that is associated with each voltage wave. In dealing with current, it is important to keep track of the *sign* of the current as it relates to the voltage waves and their polarities. Referring to Figs. 13.16*a* and 13.17, we use the convention in which current associated with a *forward-z* traveling voltage wave of positive polarity is positive. This would result in current that flows in the clockwise direction, as shown in the Fig. 13.16*a*. Current associated with a *backward-z* traveling voltage wave of positive polarity (thus flowing counterclockwise) is negative. Such a case is illustrated in Fig. 13.17. In our two-dimensional transmission line drawings, we assign positive polarity to voltage waves propagating in *either* direction if the upper conductor carries a positive charge and the lower conductor a negative charge. In Figs. 13.16*a* and 13.17, both voltage waves are of positive polarity, so their associated currents will be net positive for the forward wave, and net negative for the backward wave. In general, we write

$$I^+ = \frac{V^+}{Z_0} \tag{51}$$

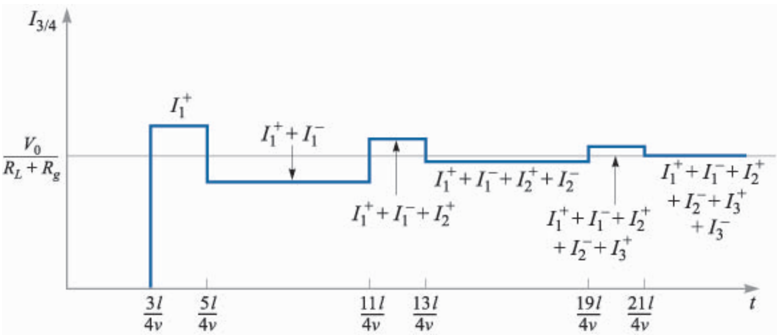
and

$$I^- = -\frac{V^-}{Z_0} \tag{52}$$

Finding the current associated with a backward-propagating voltage wave immediately requires a minus sign, as (52) indicates.



(*a*)



(*b*)

FIGURE 13.19
(*a*) Current reflection diagram for the line of Fig. 13.17 as obtained from the voltage diagram of Fig. 13.18*a*. (*b*) Current at the $z = 3l/4$ position as determined from the current reflection diagram, showing the expected steady state value of $V_0/(R_L + R_g)$.

Fig. 13.19*a* shows the current reflection diagram that is derived from the voltage diagram of Fig. 13.18*a*. Note that the current values are labeled in terms of the voltage values, with the appropriate sign added as per (51) and (52). Once the current diagram is constructed, current at a given location and time can be found in exactly the same manner as voltage is found using the voltage diagram.

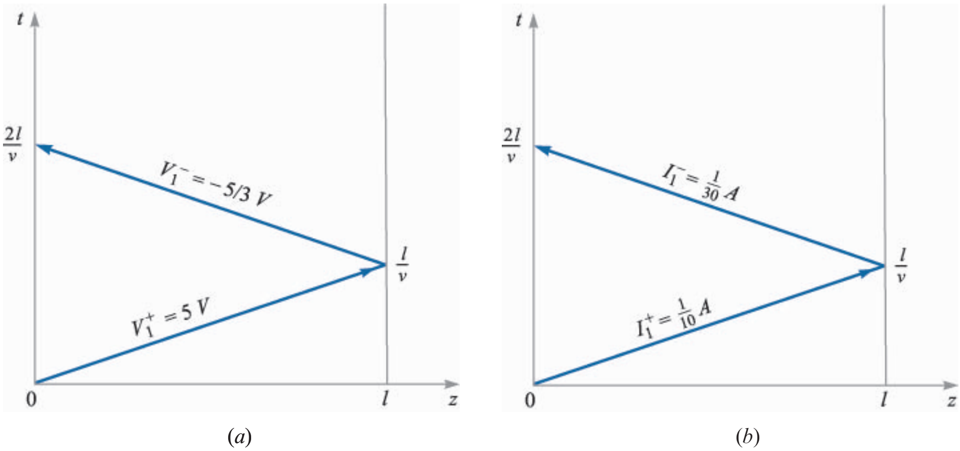


FIGURE 13.20
Voltage (*a*) and current (*b*) reflection diagrams for Example 13.5.

Fig. 13.19*b* shows the current as a function of time at the $z = (3/4)l$ position, determined by summing the current wave values as they cross the vertical line drawn at that location.

Example 13.5

In the line shown in Fig. 13.17, $R_g = Z_0 = 50\,\Omega$, $R_L = 25\,\Omega$, and the battery voltage is $V_0 = 10$ V. The switch is closed at time $t = 0$. Determine the voltage at the load resistor and the current in the battery as functions of time.

Solution. Voltage and current reflection diagrams are shown in Fig. 13.20*a* and *b*. At the moment the switch is closed, half the battery voltage appears across the 50 ohm resistor, with the other half comprising the initial voltage wave. Thus $V_1^+ = (1/2)V_0 = 5$ V. The wave reaches the 25 ohm load, where it reflects with reflection coefficient

$$\Gamma_L = \frac{25 - 50}{25 + 50} = -\frac{1}{3}$$

So $V_1^- = -(1/3)V_1^+ = -5/3$ V. This wave returns to the battery, where it encounters reflection coefficient, $\Gamma_g = 0$. Thus, no further waves appear; steady state is reached.

Once the voltage wave values are known, the current reflection diagram can be constructed. The values for the two current waves are

$$I_1^+ = \frac{V_1^+}{Z_0} = \frac{5}{50} = \frac{1}{10}\ \text{A}$$

and

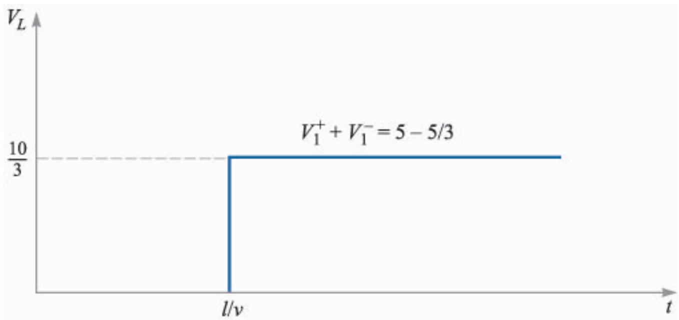
$$I_1^- = -\frac{V_1^-}{Z_0} = -\left(-\frac{5}{3}\right)\left(\frac{1}{50}\right) = \frac{1}{30} \text{ A}$$

Note that no attempt is made here to derive I_1^- from I_1^+. They are both obtained independently from their respective voltages.

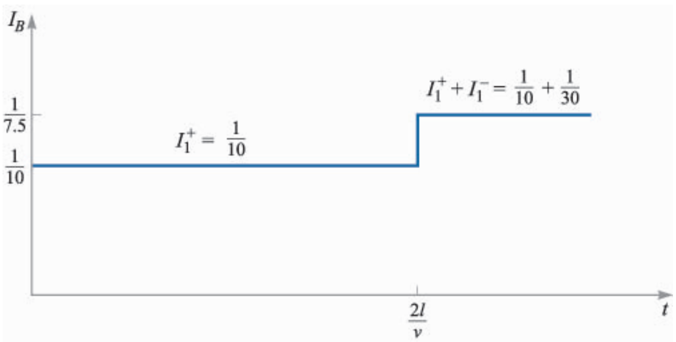
The voltage at the load as a function of time is now found by summing the voltages along the vertical line at the load position. The resulting plot is shown in Fig. 13.21*a*. Current in the battery is found by summing the currents along the vertical axis, with the resulting plot shown as Fig. 13.21*b*. Note that in steady state, we treat the circuit as lumped, with the battery in series with the 50 and 25 ohm resistors. Therefore, we expect to see a steady-state current through the battery (and everywhere else) of

$$I_B(\text{steady state}) = \frac{10}{50+25} = \frac{1}{7.5} \text{ A}$$

This value is also found from the current reflection diagram for $t > 2l/v$. Similarly, the steady-state load voltage should be



(*a*)



(*b*)

FIGURE 13.21
Voltage across the load (*a*), and current in the battery (*b*), as determined from the reflection diagrams of Fig. 13.20 (Example 13.5).

$$V_L(\text{steady state}) = V_0 \frac{R_L}{R_g + R_L} = \frac{(10)(25)}{50 + 25} = \frac{10}{3} \text{ V}$$

which is found also from the voltage reflection diagram for $t > l/v$.

Another type of transient problem involves lines that are *initially charged*. In these cases, the manner in which the line discharges through a load is of interest. Consider the situation shown in Fig. 13.22, in which a charged line of characteristic impedance Z_0 is discharged through a resistor of value R_g when a switch at the resistor location is closed.[6] We consider the resistor at the $z = 0$ location; the other end of the line is open (as would be necessary) and is located at $z = l$.

When the switch is closed, current I_R begins to flow through the resistor, and the line discharge process begins. This current does not immediately flow everywhere in the transmission line, but begins at the resistor, and establishes its presence at more distant parts of the line as time progresses. By analogy, consider a long line of automobiles at a red light. When the light turns green, the cars at the front move through the intersection first, followed successively by those further toward the rear. The point which divides cars in motion and those standing still is in fact a wave which propagates toward the back of the line. In the transmission line, the flow of charge progresses in a similar way. A voltage wave, V_1^+, is initiated and propagates to the right. To the left of its leading edge, charge is in motion; to the right of the leading edge, charge is stationary, and carries its original density. Accompanying the charge in motion to the left of V_1^+ is a drop in the charge density as the discharge process occurs, and so the line voltage to the left of V_1^+ is partially reduced. This voltage will be given by the sum of the initial voltage, V_0, and V_1^+, which means that V_1^+ must in fact be

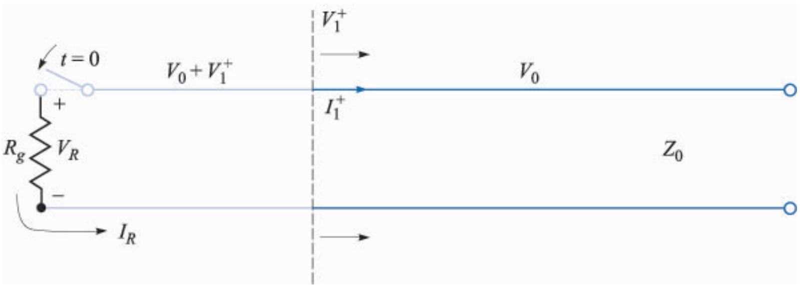


FIGURE 13.22
In an initially charged line, closing the switch as shown initiates a voltage wave of opposite polarity to that of the initial voltage. The wave thus depletes the line voltage and will fully discharge the line in one round trip if $R_g = Z_0$.

[6] Even though this is a load resistor, we will call it R_g since it is located at the front (generator) end of the line.

negative (or of opposite sign to V_0). The line discharge process is analyzed by keeping track of V_1^+ as it propagates and undergoes multiple reflections at the two ends. Voltage and current reflection diagrams are used for this purpose in much the same way as before.

Referring to Fig. 13.22, we see that for positive V_0 the current flowing through the resistor will be counterclockwise, and hence negative. We also know that continuity requires that the resistor current be equal to the current associated with the voltage wave, or

$$I_R = -I_1^+ = -\frac{V_1^+}{Z_0}$$

Now the resistor voltage will be

$$V_R = V_0 + V_1^+ = I_R R_g = -I_1^+ R_g = -\frac{V_1^+}{Z_0} R_g$$

We solve for V_1^+ to obtain

$$V_1^+ = \frac{-V_0 Z_0}{Z_0 + R_g} \tag{53}$$

Having found V_1^+, we can set up the voltage and current reflection diagrams. That for voltage is shown in Fig. 13.23. Note that the initial condition of voltage V_0 everywhere on the line is accounted for by assigning voltage V_0 to the horizontal axis of the voltage diagram. The diagram is otherwise drawn as before, but with $\Gamma_L = 1$ (at the open-circuited load end). Variations in how the line discharges thus depend on the resistor value at the switch end, R_g, which determines the reflection coefficient, Γ_g, at that location. The current reflection diagram is derived from the voltage diagram in the usual way. There is no initial current to consider.

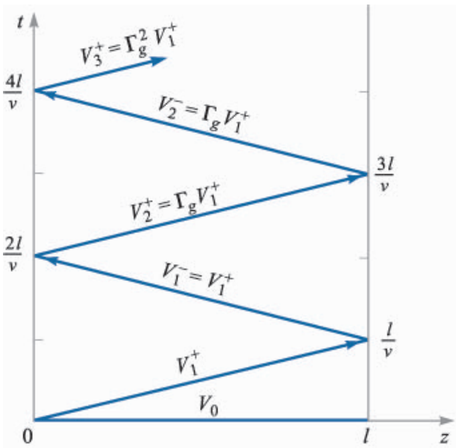


FIGURE 13.23
Voltage reflection diagram for the charged line of Fig. 13.22, showing the initial condition of V_0 everywhere on the line at $t = 0$.



FIGURE 13.24
Voltage across the resistor as a function of time, as determined from the reflection diagram of Fig. 13.23, in which $R_g = Z_0$ ($\Gamma_g = 0$).

A special case of practical importance is that in which the resistor is matched to the line, or $R_g = Z_0$. In this case, Eq. (53) gives $V_1^+ = -V_0/2$. The line fully discharges in one round-trip of V_1^+, and produces a voltage across the resistor of value $V_R = V_0/2$, which persists for time $T = 2l/v$. The resistor voltage as a function of time is shown in Fig. 13.24. The transmission line in this application is known as a *pulse-forming line*. Pulses that are generated in this way are well-formed and of low noise, provided the switch is sufficiently fast. Commercial units are available that are capable of generating high-voltage pulses of widths on the order of a few nanoseconds, using thyratron-based switches.

When the resistor is not matched to the line, full discharge still occurs, but does so over several reflections, leading to a complicated pulse shape.

Example 13.6

In the charged line of Fig. 13.22, the characteristic impedance is $Z_0 = 100\ \Omega$, and $R_g = 100/3\ \Omega$. The line is charged to an initial voltage, $V_0 = 160$ V, and the switch is closed at time $t = 0$. Determine and plot the voltage and current through the resistor for time $0 < t < 8l/v$ (four round-trips).

Solution. With the given values of R_g and Z_0, Eq. (47) gives $\Gamma_g = -1/2$. Then, with $\Gamma_L = 1$, and using (53), we find

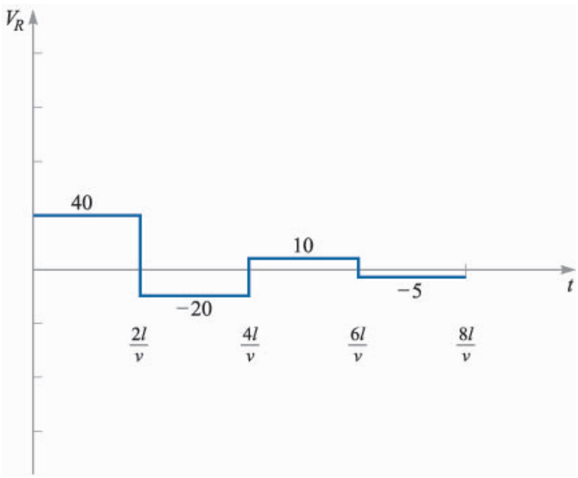
$$V_1^+ = V_1^- = -\frac{3}{4}V_0 = -120\,\text{V}$$
$$V_2^+ = V_2^- = \Gamma_g V_1^- = +60\,\text{V}$$
$$V_3^+ = V_3^- = \Gamma_g V_2^- = -30\,\text{V}$$
$$V_4^+ = V_4^- = \Gamma_g V_3^- = +15\,\text{V}$$

Using these values on the voltage reflection diagram, we evaluate the voltage in time at the resistor location by moving up the left-hand vertical axis, adding voltages as we progress, and beginning with $V_0 + V_1^+$ at $t = 0$. Note that when we add voltages along

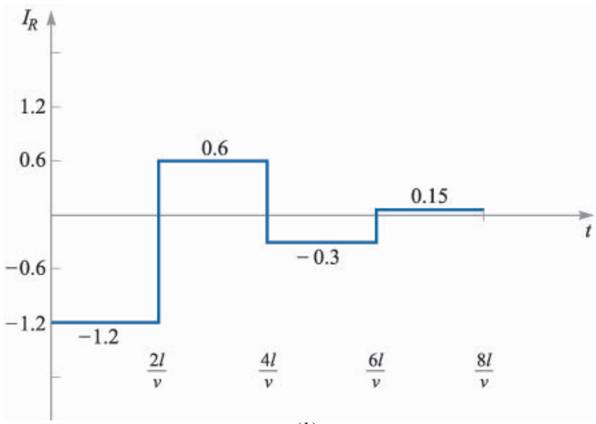
the vertical axis, we are encountering the intersection points between incident and reflected waves, which occur (in time) at each integer multiple of $2l/v$. So, when moving up the axis, we add the voltages of *both* waves to our total at each occurrence. The voltage within each time interval is thus:

$$
\begin{aligned}
V_R &= V_0 + V_1^+ = 40\text{ V} && (0 < t < 2l/v)\\
&= V_0 + V_1^+ + V_1^- + V_2^+ = -20\text{ V} && (2l/v < t < 4l/v)\\
&= V_0 + V_1^+ + V_1^- + V_2^+ + V_2^- + V_3^+ = 10\text{ V} && (4l/v < t < 6l/v)\\
&= V_0 + V_1^+ + V_1^- + V_2^+ + V_2^- + V_3^+ + V_3^- + V_4^+ = -5\text{ V} && (6l/v < t < 8l/v)
\end{aligned}
$$

The resulting voltage plot over the desired time range is shown in Fig. 13.25*a*.



(*a*)



(*b*)

FIGURE 13.25
Resistor voltage (*a*) and current (*b*) as functions of time for the line of Fig. 13.22, with values as specified in Example 13.6.

The current through the resistor is most easily obtained by dividing the voltages in Fig. 13.25*a* by $-R_g$. As a demonstration, we can also use the current diagram of Fig. 13.19*a* to obtain this result. Using (51) and (52), we evaluate the current waves as follows:

$$\begin{aligned}
I_1^+ &= V_1^+/Z_0 = -1.2\,\text{A} \\
I_1^- &= -V_1^-/Z_0 = +1.2\,\text{A} \\
I_2^+ &= -I_2^- = V_2^+/Z_0 = +0.6\,\text{A} \\
I_3^+ &= -I_3^- = V_3^+/Z_0 = -0.30\,\text{A} \\
I_4^+ &= -I_4^- = V_4^+/Z_0 = +0.15\,\text{A}
\end{aligned}$$

Using the above values on the current reflection diagram, Fig. 13.19*a*, we add up currents in the resistor in time by moving up the left-hand axis, as we did with the voltage diagram. The result is shown in Fig. 13.25*b*. As a further check to the correctness of our diagram construction, we note that current at the open end of the line ($Z = l$) must always be zero. Therefore, summing currents up the right-hand axis must give a zero result for all time. The reader is encouraged to verify this.

SUGGESTED REFERENCES

1. Brown, R. G., R. A. Sharpe, W. L. Hughes, and R. E. Post: "Lines, Waves, and Antennas," 2d ed., The Ronald Press Company, New York, 1973. Transmission lines are covered in the first six chapters, with numerous examples.
2. Cheng, D. K.: "Field and Wave Electromagnetics," 2nd ed., Addison-Wesley Publishing Company, Reading, Mass., 1989. Provides numerous examples of Smith Chart problems and transients.
3. Seshadri, S. R.: "Fundamentals of Transmission Lines and Electromagnetic Fields," Addison-Wesley Publishing Company, Reading, Mass., 1971.
4. Weeks, W. L.: "Transmission and Distribution of Electrical Energy," Harper and Row, Publishers, New York, 1981. Line parameters for various configurations of power transmission and distribution systems are discussed in Chap. 2, along with typical parameter values.

PROBLEMS

13.1 The parameters of a certain transmission line operating at 6×10^8 rad/s are $L = 0.4\,\mu$H/m, $C = 40$ pF/m, $G = 80$ mS/m, and $R = 20\ \Omega$/m. (a) Find γ, α, β, λ, and Z_0. (b) If a voltage wave travels 20 m down the line, by what percentage is its amplitude reduced, and by how many degrees is its phase shifted?

13.2 A lossless transmission line with $Z_0 = 60\,\Omega$ is being operated at 60 MHz. The velocity on the line is 3×10^8 m/s. If the line is short-circuited at $z = 0$, find Z_{in} at $z =$: (a) -1 m; (b) -2 m; (c) -2.5 m; (d) -1.25 m.

13.3 The characteristic impedance of a certain lossless transmission line is $72\,\Omega$. If $L = 0.5\,\mu$H/m, find: (a) C; (b) v_p; (c) β if $f = 80$ MHz. (d) The line is terminated with a load of 60 Ω. Find Γ and s.

13.4 A lossless transmission line having $Z_0 = 120\ \Omega$ is operating at $\omega = 5 \times 10^8$ rad/s. If the velocity on the line is 2.4×10^8 m/s, find: (a) L; (b) C. (c) Let Z_L be represented by an inductance of 0.6 μH in series with a 100-Ω resistance. Find Γ and s.

13.5 Two characteristics of a certain lossless transmission line are $Z_0 = 50\,\Omega$ and $\gamma = 0 + j0.2\pi\,\text{m}^{-1}$ at $f = 60$ MHz: (a) find L and C for the line. (b) A load $Z_L = 60 + j80\ \Omega$ is located at $z = 0$. What is the shortest distance from the load to a point at which $Z_{in} = R_{in} + j0$?

13.6 The propagation constant of a lossy transmission line is $1 + j2\,\text{m}^{-1}$, and its characteristic impedance is $20 + j0\ \Omega$ at $\omega = 1$ Mrad/s. Find L, C, R, and G for the line.

13.7 The dimensions of the outer conductor of a coaxial cable are b and c, $c > b$. Assume $\sigma = \sigma_c$ and let $\mu = \mu_0$. Find the magnetic energy stored per unit length in the region $b < r < c$ for a uniformly distributed total current I flowing in the opposite directions in the inner and outer conductors.

13.8 The conductors of a coaxial transmission line are copper ($\sigma_c = 5.8 \times 10^7$ S/m), and the dielectric is polyethylene ($\epsilon'_R = 2.26$, $\sigma/\omega\epsilon' = 0.0002$). If the inner radius of the outer conductor is 4 mm, find the radius of the inner conductor so that: (a) $Z_0 = 50\ \Omega$; (b) $C = 100$ pF/m; (c) $L = 0.2\,\mu$H/m. A lossless line can be assumed.

13.9 Two aluminum-clad steel conductors are used to construct a two-wire transmission line. Let $\sigma_{\text{Al}} = 3.8 \times 10^7$ S/m, $\sigma_{\text{St}} = 5 \times 10^6$ S/m, and $\mu_{\text{St}} = 100\,\mu$H/m. The radius of the steel wire is 0.5 in., and the aluminum coating is 0.05 in thick. The dielectric is air, and the center-to-center wire separation is 4 in. Find C, L, G, and R for the line at 10 MHz.

13.10 Each conductor of a two-wire transmission line has a radius of 0.5 mm; their center-to-center separation is 0.8 cm. Let $f = 150$ MHz, and assume σ and σ_c are zero. Find the dielectric constant of the insulating medium if: (a) $Z_0 = 300\ \Omega$; (b) $C = 20$ pF/m; (c) $v_p = 2.6 \times 10^8$ m/s.

13.11 Pertinent dimensions for the transmission line shown in Fig. 13.4 are $b = 3$ mm and $d = 0.2$ mm. The conductors and the dielectric are nonmagnetic. (a) If the characteristic impedance of the line is 15 Ω, find ϵ'_R. Assume a low-loss dielectric. (b) Assume copper conductors and operation at 2×10^8 rad/s. If $RC = GL$, determine the loss tangent of the dielectric.

13.12 A transmission line constructed from perfect conductors and an air dielectric is to have a maximum dimension of 8 mm for its cross section. The line is to be used at high frequencies. Specify the dimensions if it is: (a) a two-wire line with $Z_0 = 300\ \Omega$; (b) a planar line with $Z_0 = 15\ \Omega$; (c) a 72-Ω coax having a zero-thickness outer conductor.

13.13 The incident voltage wave on a certain lossless transmission line for which $Z_0 = 50\ \Omega$ and $v_p = 2 \times 10^8$ m/s is $V^+(z, t) = 200\cos(\omega t - \pi z)$ V. (a) Find ω. (b) Find $I^+(z, t)$. The section of line for which $z > 0$ is replaced by a load $Z_L = 50 + j30\ \Omega$ at $z = 0$. Find: (c) Γ_L; (d) $V_s^-(z)$; (e) V_s at $z = -2.2$ m.

13.14 Coaxial lines 1 and 2 have the following parameters: $\mu_1 = \mu_2 = \mu_0$, $\sigma_1 = \sigma_2 = 0$, $\epsilon'_{R1} = 2.25$, $\epsilon'_{R2} = 4$, $a_1 = a_2 = 0.8$ mm, $b_1 = 6$ mm, $b_2 = 3$ mm, $Z_{L2} = Z_{02}$, and Z_{L1} is Z_{in2}. (a) Find Z_{01} and Z_{02}. (b) Find s on line 1. (c) If a 20-cm length of line 1 is inserted immediately in front of Z_{L2} and $f = 300$ MHz, find s on line 2.

13.15 For the transmission line represented in Fig. 13.26, find $V_{s,out}$ if $f =$:(a) 60 Hz; (b) 500 kHz.

13.16 A 300-Ω transmission line is 0.8 m long and terminated with a short circuit. The line is operating in air with a wavelength of 0.8 m and is lossless. (a) If the input voltage amplitude is 10 V, what is the maximum voltage amplitude at any point on the line? (b) What is the current amplitude in the short circuit?

13.17 Determine the average power absorbed by each resistor in Fig. 13.27.

13.18 The line shown in Fig. 13.28 is lossless. Find s on both sections 1 and 2.

13.19 A lossless transmission line is 50 cm in length and operating at a frequency of 100 MHz. The line parameters are $L = 0.2\ \mu$H/m and $C = 80$ pF/m. The line is terminated in a short circuit at $z = 0$, and there is a load $Z_L = 50 + j20\ \Omega$ across the line at location $z = -20$ cm. What average power is delivered to Z_L if the input voltage is $100\angle 0°$ V?

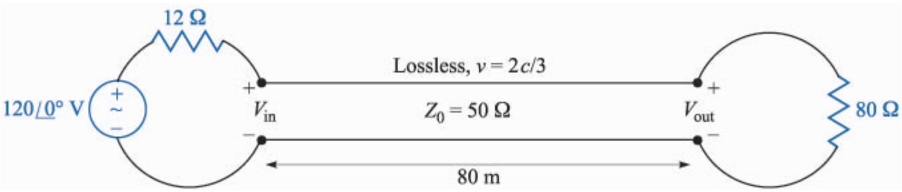


FIGURE 13.26
See Problem 15.

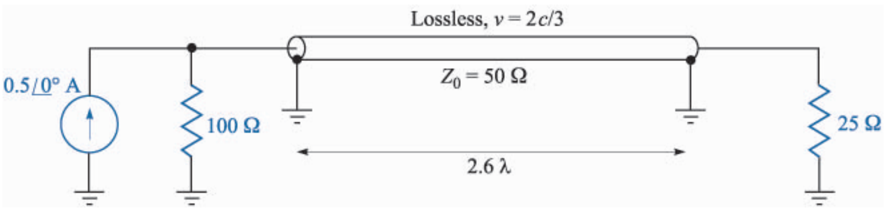


FIGURE 13.27
See Problem 17.

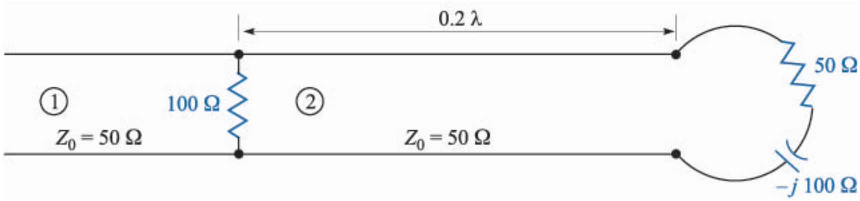


FIGURE 13.28
See Problem 18.

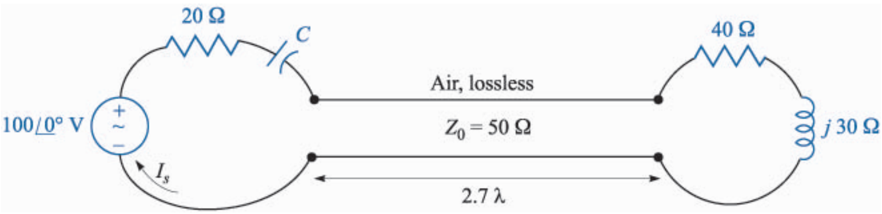


FIGURE 13.29
See Problem 20.

13.20 (a) Determine s on the transmission line of Fig. 13.29. Note that the dielectric is air. (b) Find the input impedance. (c) If $1/\omega C = 10\ \Omega$, find I_s. (d) What value of C will produce a maximum value for $|I_s|$ at $\omega = 1$ Grad/s? For this value of C, calculate the average power: (e) supplied by the source; (f) delivered to $Z_L = 40 + j30\ \Omega$.

13.21 A lossless line having an air dielectric has a characteristic impedance of 400 Ω. The line is operating at 200 MHz and $Z_{in} = 200 - j200\ \Omega$. Use analytic methods or the Smith chart (or both) to find: (a) s, (b) Z_L, if the line is 1 m long; (c) the distance from the load to the nearest voltage maximum.

13.22 A lossless two-wire line has a characteristic impedance of 300 Ω and a capacitance of 15 pF/m. The load at $z = 0$ consists of a 600-Ω resistor in parallel with a 10-pF capacitor. If $\omega = 10^8$ rad/s and the line is 20 m long, use the Smith chart to find: (a) $|\Gamma_L|$; (b) s; (c) Z_{in}.

13.23 The normalized load on a lossless transmission line is $2 + j1$. Let $l = 20$ m and make use of the Smith chart to find: (a) the shortest distance from the load to a point at which $z_{in} = r_{in} + j0$, where $r_{in} > 0$; (b) z_{in} at this point. (c) The line is cut at this point and the portion containing z_L is thrown away. A resistor $r = r_{in}$ of part (a) is connected across the line. What is s on the remainder of the line? (d) What is the shortest distance from this resistor to a point at which $z_{in} = 2 + j1$?

13.24 With the aid of the Smith chart, plot a curve of $|Z_{in}|$ vs. l for the transmission line shown in Fig. 13.30. Cover the range $0 < l/\lambda < 0.25$.

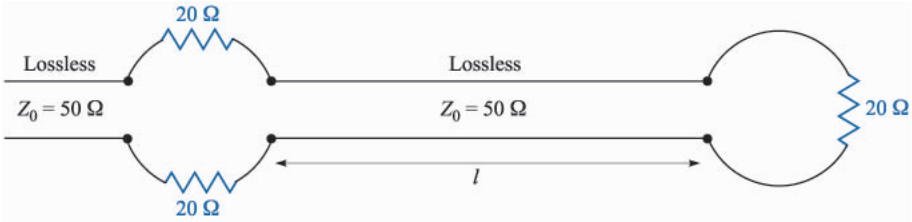


FIGURE 13.30
See Problem 24.

13.25 A 300-Ω transmission line is short-circuited at $z = 0$. A voltage maximum, $|V|_{max} = 10$ V, is found at $z = -25$ cm, and the minimum voltage, $|V|_{min} = 0$ is at $z = -50$ cm. Use the Smith chart to find Z_L (with the short circuit replaced by the load) if the voltage readings are: (a) $|V|_{max} = 12$ V at $z = -5$ cm, and $|V|_{min} = 5$ V; (b) $|V|_{max} = 17$ V at $z = -20$ cm, and $|V|_{min} = 0$.

13.26 A lossless 50-Ω transmission line operates with a velocity that is 3/4 c. A load $Z_L = 60 + j30\ \Omega$ is located at $z = 0$. Use the Smith chart to find: (a) s; (b) the distance from the load to the nearest voltage minimum if $f = 300$ MHz; (c) the input impedance if $f = 200$ MHz and the input is at $z = -110$ cm.

13.27 The characteristic admittance ($Y_0 = 1/Z_0$) of a lossless transmission line is 20 mS. The line is terminated in a load $Y_L = 40 - j20$ mS. Make use of the Smith chart to find: (a) s; (b) Y_{in} if $l = 0.15\lambda$; (c) the distance in wavelengths from Y_L to the nearest voltage maximum.

13.28 The wavelength on a certain lossless line is 10 cm. If the normalized input impedance is $z_{in} = 1 + j2$, use the Smith chart to determine: (a) s; (b) z_L, if the length of the line is 12 cm; (c) x_L, if $z_L = 2 + jx_L$ where $x_L > 0$.

13.29 A standing wave ratio of 2.5 exists on a lossless 60-Ω line. Probe measurements locate a voltage minimum on the line whose location is marked by a small scratch on the line. When the load is replaced by a short circuit, the minima are 25 cm apart, and one minimum is located at a point 7 cm toward the source from the scratch. Find Z_L.

13.30 A 2-wire line constructed of lossless wire of circular cross section is gradually flared into a coupling loop that looks like an egg beater. At the point X, indicated by the arrow in Fig. 13.31, a short circuit is placed across the line. A probe is moved along the line and indicates that the first voltage minimum to the left of X is 16 cm from X. With the short circuit removed, a voltage minimum is found 5 cm to the left of X, and a voltage maximum is located that is 3 times the voltage of the minimum. Use the Smith chart to determine: (a) f; (b) s; (c) the normalized input impedance of the egg beater as seen looking to the right at point X.

13.31 In order to compare the relative sharpness of the maxima and minima of a standing wave, assume a load $z_L = 4 + j0$ is located at $z = 0$. Let

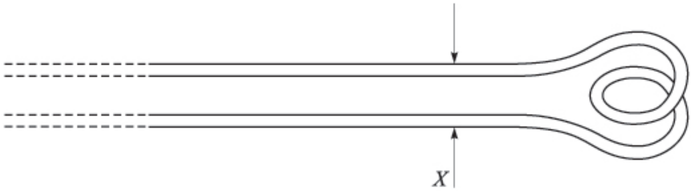


FIGURE 13.31
See Problem 30.

$|V|_{min} = 1$ and $\lambda = 1$ m. Determine the width of the: (a) minimum where $|V| < 1.1$; (b) maximum where $|V| > 4/1.1$.

13.32 A lossless line is operating with $Z_0 = 40\ \Omega$, $f = 20$ MHz, and $\beta = 7.5\pi$ rad/m. With a short circuit replacing the load, a minimum is found at a point on the line marked by a small spot of puce paint. With the load installed, it is found that $s = 1.5$ and a voltage minimum is located 1 m toward the source from the puce dot. (a) Find Z_L. (b) What load would produce $s = 1.5$ with $|V|_{max}$ at the paint spot?

13.33 In Fig. 13.14, let $Z_L = 40 - j10\ \Omega$, $Z_0 = 50\ \Omega$, $f = 800$ MHz, and $v = c$. (a) Find the shortest length d_1 of a short-circuited stub, and the shortest distance d that it may be located from the load to provide a perfect match on the main line to the left of the stub. (b) Repeat for an open-circuited stub.

13.34 The lossless line shown in Fig. 13.32 is operating with $\lambda = 100$ cm. If $d_1 = 10$ cm, $d = 25$ cm, and the line is matched to the left of the stub, what is Z_L?

13.35 A load, $Z_L = 25 + j75\ \Omega$, is located at $z = 0$ on a lossless two-wire line for which $Z_0 = 50\ \Omega$ and $v = c$. (a) If $f = 300$ MHz, find the shortest distance d ($z = -d$) at which the input admittance has a real part equal to $1/Z_0$ and a negative imaginary part. (b) What value of capacitance C should be connected across the line at that point to provide unity standing wave ratio on the remaining portion of the line?

13.36 The two-wire lines shown in Fig. 13.33 are all lossless and have $Z_0 = 200\ \Omega$. Find d and the shortest possible value for d_1 to provide a matched load if $\lambda = 100$ cm.

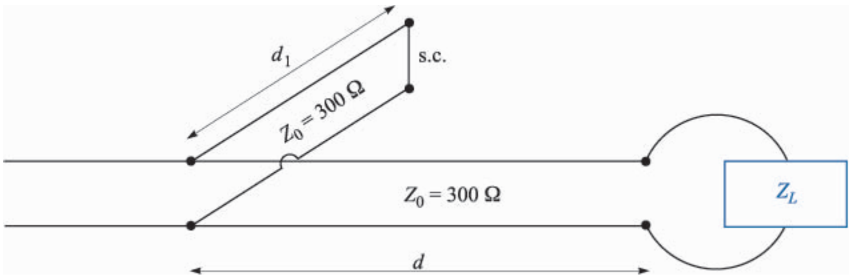


FIGURE 13.32
See Problem 34.

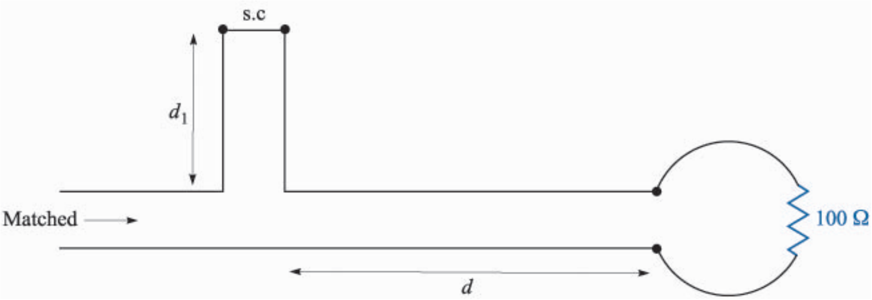


FIGURE 13.33
See Problem 36.

13.37 In the transmission line of Fig. 13.17, $R_L = Z_0 = 50\ \Omega$, and $R_g = 25\ \Omega$. Determine and plot the voltage at the load resistor and the current in the battery as functions of time by constructing appropriate voltage and current reflection diagrams.

13.38 Repeat Problem 37, with $Z_0 = 50\ \Omega$, and $R_L = R_g = 25\ \Omega$. Carry out the analysis for the time period $0 < t < 8l/v$.

13.39 In the transmission line of Fig. 13.17, $Z_0 = 50\ \Omega$, and $R_L = R_g = 25\ \Omega$. The switch is closed at $t = 0$ and *is opened again* at time $t = l/4v$, thus creating a rectangular voltage *pulse* in the line. Construct an appropriate voltage reflection diagram for this case and use it to make a plot of the voltage at the load resistor as a function of time for $0 < t < 8l/v$ (note that the effect of opening the switch is to initiate a second voltage wave, whose value is such that it leaves a net current of zero in its wake).

13.40 In the charged line of Fig. 13.22, the characteristic impedance is $Z_0 = 100\ \Omega$, and $R_g = 300\ \Omega$. The line is charged to initial voltage, $V_0 = 160$ V, and the switch is closed at $t = 0$. Determine and plot the voltage and current through the resistor for time $0 < t < 8l/v$ (four round-trips). This problem accompanies Example 13.6 as the other special case of the basic charged line problem, in which now $R_g > Z_0$.

13.41 In the transmission line of Fig. 13.34, the switch is located *midway* down the line, and is closed at $t = 0$. Construct a voltage reflection diagram for this case, where $R_L = Z_0$. Plot the load resistor voltage as a function of time.

13.42 A simple *frozen wave generator* is shown in Fig. 13.35. Both switches are closed simultaneously at $t = 0$. Construct an appropriate voltage reflection diagram for the case in which $R_L = Z_0$. Determine and plot the load resistor voltage as a function of time.

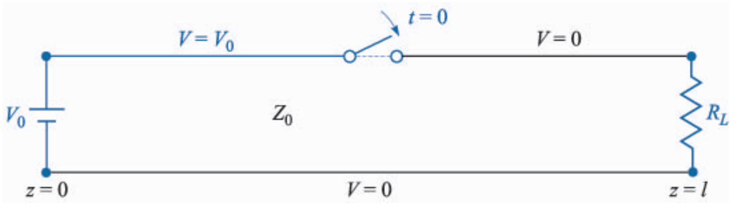


FIGURE 13.34
See Problem 41.

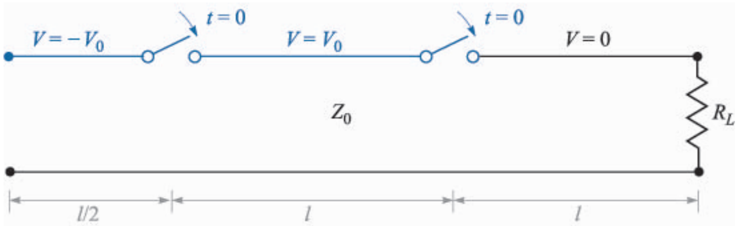


FIGURE 13.35
See Problem 42.

CHAPTER

14

WAVEGUIDE AND ANTENNA FUNDAMENTALS

As a conclusion to our study of electromagnetics, we investigate the basic principles of two important classes of devices: waveguides and antennas. In broad definitions, a waveguide is a structure through which electromagnetic waves can be transmitted from point to point, and within which the fields are confined to a certain extent. An antenna is any device that radiates electromagnetic fields into space, where the fields originate from a source that feeds the antenna through a transmission line or waveguide. The antenna thus serves as an interface between the confining line and space when used as a transmitter—or between space and the line when used as a receiver.

In our study of waveguides, we will first take a broad view of waveguide devices, to obtain a physical understanding of how they work and the conditions under which they are used. We will next explore the simple parallel-plate waveguide and study the concept of waveguide modes and the conditions under which these will occur. We will study the electric and magnetic field configurations of the guided modes using simple plane wave models and through use of the wave equation. We will then study more complicated structures, including the rectangular waveguide and the dielectric slab guide.

Our study of antennas will include the derivation of the radiated fields from an elemental dipole, beginning with the retarded vector potentials that we studied in Chap. 10. We will address issues that include the efficiency of power radiation from an antenna, and the parameters that govern this.

14.1 BASIC WAVEGUIDE OPERATION

Waveguides assume many different forms that depend on the purpose of the guide, and on the frequency of the waves to be transmitted. The simplest form (in terms of analysis) is the parallel-plate guide shown in Fig. 14.1. Other forms are the hollow-pipe guides, including the rectangular waveguide of Fig. 14.2, and the cylindrical guide, shown in Fig. 14.3. Dielectric waveguides, used primarily at optical frequencies, include the slab waveguide of Fig. 14.4 and the optical fiber, shown in Fig. 14.5. Each of these structures possesses certain advantages over the others, depending on the application and the frequency of the waves to be transmitted. All guides, however, exhibit the same basic operating principles, which we will explore in this section.

To develop an understanding of waveguide behavior, we consider the parallel-plate waveguide of Fig. 14.1. At first, we recognize this as one of the transmission line structures that we investigated in Chap. 13. So the first question that arises is: How does a waveguide differ from a transmission line to begin with? The difference lies in the form of the electric and magnetic fields within the line. To see this, consider Fig. 14.6*a*, which shows the fields when the line operates as a transmission line. A sinusoidal voltage wave, with voltage applied between conductors, leads to an electric field that is directed vertically between the conductors as shown. Since current flows only in the z direction, magnetic field will be oriented in and out of the page (in the y direction). The interior fields comprise a plane electromagnetic wave which propagates in the z direction (as the Poynting vector will show), since both fields lie in the transverse plane. We refer to this as a transmission line wave, which, as discussed in Chap. 13, is a transverse electromagnetic (TEM) wave. The wavevector **k**, shown in the figure, indicates the direction of wave travel, as well as the direction of power flow. With perfectly conducting plates, the electric field between plates is found by solving Eq. (29), Chap. 11, leading to Eq. (31) in that chapter.

As the frequency is increased, a remarkable change occurs in the way the fields progagate down the line. Although the original field configuration of Fig. 14.6*a* may still be present, another possibility emerges which is shown in Fig.

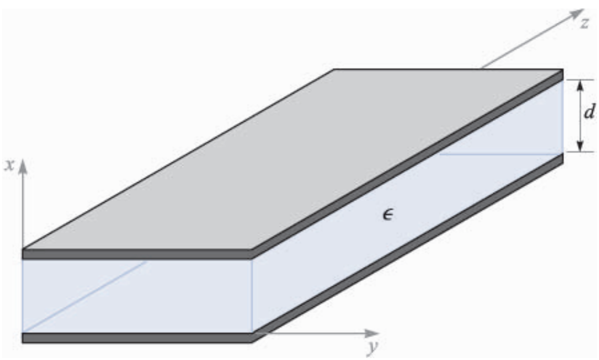


FIGURE 14.1
Parallel-plate waveguide, with metal plates at $x = 0, d$. Between the plates is a dielectric of permittivity ϵ.

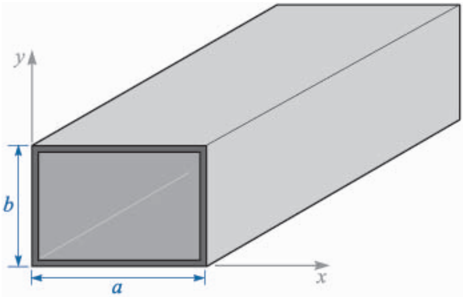


FIGURE 14.2
Rectangular waveguide.

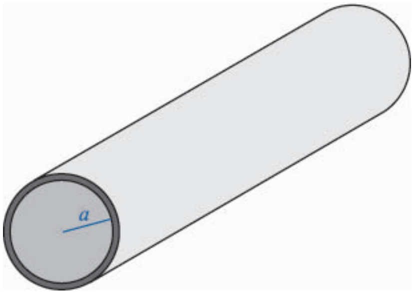


FIGURE 14.3
Cylindrical waveguide.

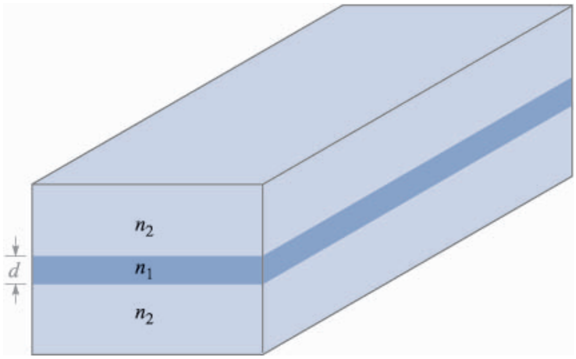


FIGURE 14.4
Symmetric dielectric slab waveguide, with slab region (refractive index n_1) surrounded by two dielectrics of index $n_2 < n_1$.

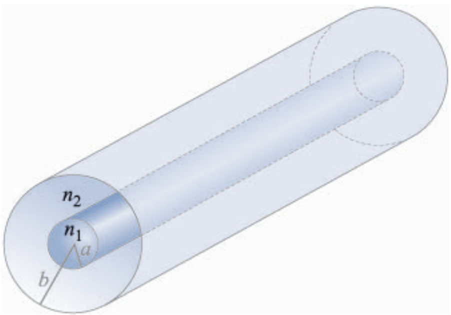


FIGURE 14.5
Optical fiber waveguide, with the core dielectric ($r < a$) of refractive index n_1. The cladding dielectric ($a < r < b$) is of index $n_2 < n_1$.

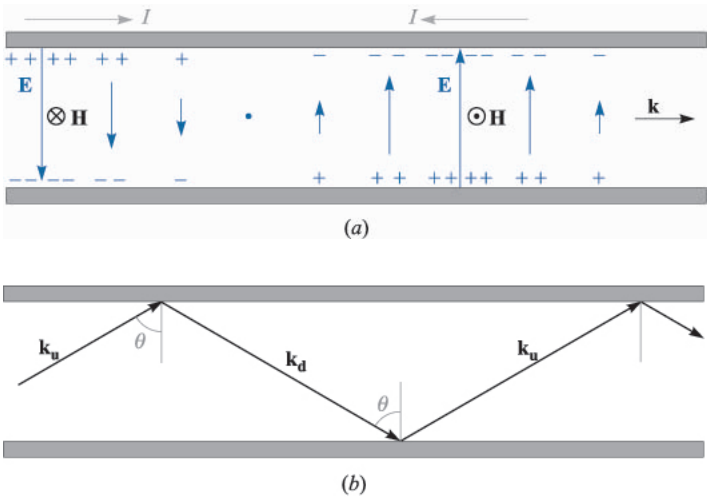


FIGURE 14.6
(*a*) Electric and magnetic fields of a TEM (transmission line) mode in a parallel-plate waveguide, forming a plane wave that propagates down the guide axis. (*b*) Plane waves that reflect from the conducting walls can produce a waveguide mode that is no longer TEM.

14.6*b*. Again, a plane wave is guided in the z direction, but does so by means of a progression of zig-zag reflections at the upper and lower plates. Wavevectors $\mathbf{k_u}$ and $\mathbf{k_d}$ are associated with the upward and downward-propagating waves, respectively, and these have identical magnitudes,

$$|\mathbf{k_u}| = |\mathbf{k_d}| = k = \omega\sqrt{\mu\epsilon}$$

For such a wave to propagate, all upward-propagating waves must be *in phase* (as must be true of all downward-propagating waves). This condition can only be satisfied at certain discrete angles of incidence, shown as θ in the figure. An allowed value of θ, along with the resulting field configuration, comprise a *waveguide mode* of the structure. Associated with each guided mode is a *cutoff frequency*. If the operating frequency is below the cutoff frequency, the mode will not propagate. If above cutoff, the mode propagates. The TEM mode, however, has no cutoff; it will be supported at any frequency. At a given frequency, the guide may support several modes, the quantity of which depends on the plate separation and on the dielectric constant of the interior medium, as will be shown. The number of modes increases as the frequency is raised.

So to answer our initial question on the distinction between transmission lines and waveguides, we can state the following: Transmission lines consist of two or more conductors and as a rule will support TEM waves (or something which could approximate such a wave). A waveguide may consist of one or more conductors, or no conductors at all, and will support waveguide modes, of forms similar to those described above. Waveguides may or may not support TEM waves, depending on the design.

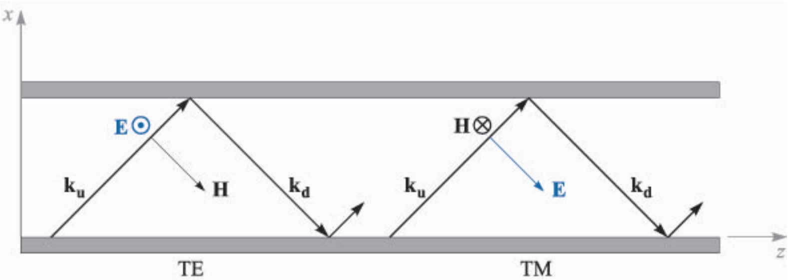


FIGURE 14.7
Plane wave representation of TE and TM modes in a parallel-plate guide.

In the parallel-plate guide, two types of waveguide modes can be supported. These are shown in Fig. 14.7 as arising from the s and p orientations of the plane wave polarizations. In a manner consistent with our previous discussions on oblique reflection (Sec. 12.5), we identify a *transverse electric* (*TE*) mode when **E** is perpendicular to the plane of incidence (s polarized); this positions **E** parallel to the transverse plane of the waveguide, as well as to the boundaries. Similarly, a *transverse magnetic* (*TM*) mode results with a p polarized wave; the entire **H** field is in the y direction and is thus within the transverse plane of the guide. Both possibilities are illustrated in the figure. Note, for example, that with **E** in the y direction (TE mode), **H** will have x and z components. Likewise, a TM mode will have x and z components of **E**.[1] In any event, the reader can verify from the geometry of Fig. 14.7 that it is not possible to achieve a purely TEM mode for values of θ other than 90°. Other wave polarizations are possible that lie between the TE and TM cases, but these can always be expressed as superpositions of TE and TM modes.

14.2 PLANE WAVE ANALYSIS OF THE PARALLEL-PLATE WAVEGUIDE

Let us now investigate the conditions under which waveguide modes will occur, using our plane wave model for the mode fields. In Fig. 14.8*a*, a zig-zag path is again shown, but this time phase fronts are drawn that are associated with two of the upward-propagating waves. The first wave has reflected twice (at the top and bottom surfaces) to form the second wave (the downward-propagating phase fronts are not shown). Note that the phase fronts of the second wave do not coincide with those of the first wave, and so the two waves are out of phase. In Fig. 14.8*b*, the wave angle has been adjusted so that the two waves are now in phase. Having satisfied this condition for the two waves, we will find that *all*

[1] Other types of modes can exist in other structures (not the parallel-plate guide) in which *both* **E** and **H** have z components. These are known as *hybrid* modes, and typically occur in guides with cylindrical cross sections, such as the optical fiber.

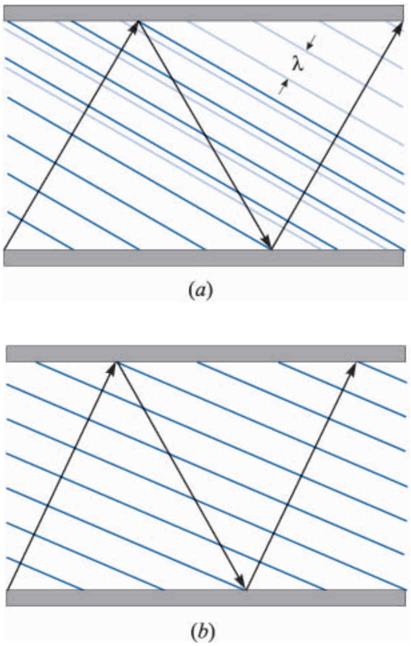


FIGURE 14.8
(*a*) Plane wave propagation in a parallel-plate guide in which the wave angle is such that the upward-propagating waves are not in phase. (*b*) The wave angle has been adjusted so that upward waves are in phase, resulting in a guided mode.

upward-propagating waves will have coincident phase fronts. The same condition will automatically occur for all downward-propagating waves. This is the requirement to establish a guided mode.

In Fig. 14.9 we show the wavevector, $\mathbf{k_u}$, and its components, along with a series of phase fronts. A drawing of this kind for $\mathbf{k_d}$ would be the same, except the x component, κ_m, would be reversed. In Sec. 12.4, we measured the phase shift per unit distance along the x and z directions by the components, k_x and k_z, which varied continuously as the direction of $\mathbf{k}$ changed. In our discussion of waveguides, we introduce a different notation, where κ_m and β_m are used for k_x and k_z. The subscript m is an integer, indicating the *mode number*. This provides a subtle hint that β_m and κ_m will assume only certain discrete values that correspond to certain allowed directions of $\mathbf{k_u}$ and $\mathbf{k_d}$, such that our coincident phase front requirement is satisfied.[2] From the geometry we see that for any value of m,

$$\boxed{\beta_m = \sqrt{k^2 - \kappa_m^2}} \tag{1}$$

Use of the symbol β_m for the z components of $\mathbf{k_u}$ and $\mathbf{k_d}$ is appropriate because β_m will ultimately be the phase constant for the mth waveguide mode, measuring

[2] Subscripts (m) are not shown on $\mathbf{k_u}$ and $\mathbf{k_d}$, but are understood. Changing m does not affect the magnitudes of these vectors—only their directions.

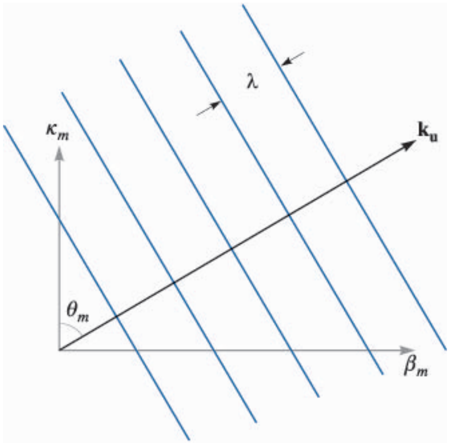


FIGURE 14.9
The components of the upward wavevector are κ_m and β_m, the transverse and axial phase constants. To form the downward wavevector, $\mathbf{k_d}$, the direction of κ_m is reversed.

phase shift per distance down the guide; it is also used to determine the phase velocity of the mode, ω/β_m, and the group velocity, $d\omega/d\beta_m$.

Throughout our discussion, we will assume that the medium within the guide is lossless and nonmagnetic, so that

$$k = \omega\sqrt{\mu_0\epsilon'} = \frac{\omega\sqrt{\epsilon'_R}}{c} = \frac{\omega n}{c} \tag{2}$$

which we express either in terms of the dielectric constant, ϵ'_R, or the refractive index, n, of the medium.

It is κ_m, the x component of $\mathbf{k_u}$ and $\mathbf{k_d}$, that will be useful to us in quantifying our requirement on coincident phase fronts through a condition known as *transverse resonance*. This condition states that the net phase shift measured during a round-trip over the full transverse dimension of the guide must be an integer multiple of 2π radians. This is another way of stating that all upward (or downward) propagating plane waves must have coincident phases. The various segments of this round-trip are illustrated in Fig. 14.10. We assume for this exercise that the waves are frozen in time, and that an observer moves vertically over the round-trip, measuring phase shift along the way. In the first segment (Fig. 14.10*a*) the observer starts at a position just above the lower conductor and moves vertically to the top conductor through distance d. The measured phase shift over this distance is $\kappa_m d$ rad. On reaching the top surface, the observer will note a possible phase shift on reflection (Fig. 14.10*b*). This will be π if the wave is TE polarized, and will be zero if the wave is TM polarized (see Fig. 14.11 for a demonstration of this). Next, the observer moves along the reflected wave phases down to the lower conductor and again measures a phase shift of $\kappa_m d$ (Fig. 14.10*c*). Finally, after including the phase shift on reflection at the bottom conductor, the observer is back at the original starting point, and is noting the phase of the next upward-propagating wave.

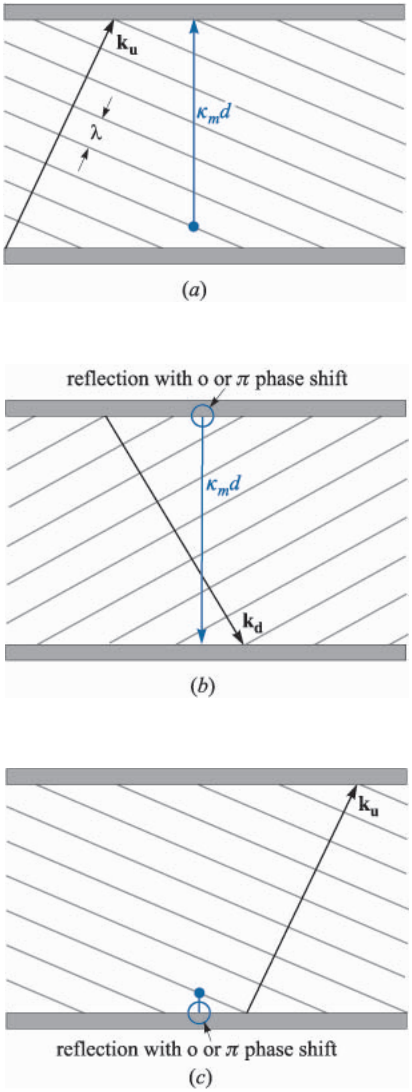


FIGURE 14.10
The net phase shift over a round-trip in the parallel-plate guide is found by first measuring the transverse phase shift between plates of the initial upward wave (*a*); next, the transverse phase shift in the reflected (downward) wave is measured, while accounting for the reflective phase shift at the top plate (*b*); finally, the phase shift on reflection at the bottom plate is added, thus returning to the starting position, but with a new upward wave (*c*). Transverse resonance occurs if the phase at the final point is the same as that at the starting point (the two upward waves are in phase)

The total phase shift over the round-trip is required to be an integer multiple of 2π:

$$\boxed{\kappa_m d + \phi + \kappa_m d + \phi = 2m\pi} \tag{3}$$

where ϕ is the phase shift on reflection at each boundary. Note that with $\phi = \pi$ (TE waves) or 0 (TM waves) the net reflective phase shift over a round-trip is 2π

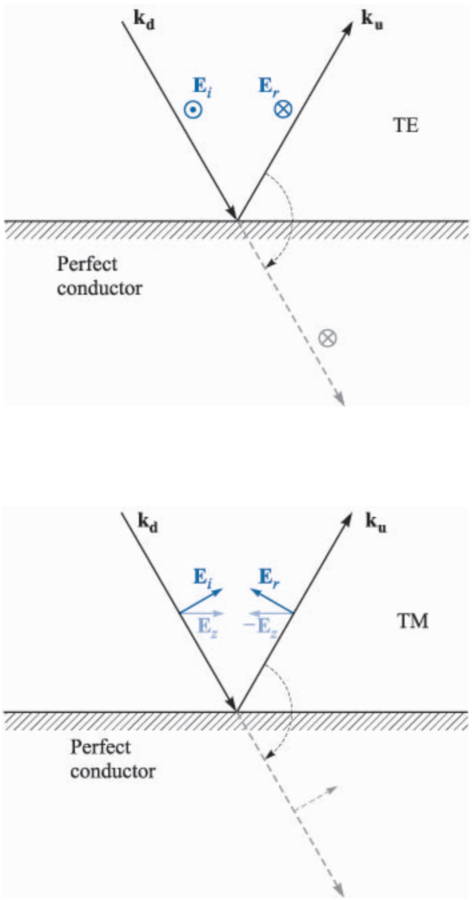


FIGURE 14.11
The phase shift of a wave on reflection from a perfectly conducting surface depends on whether the incident wave is TE (*s*-polarized) or TM (*p*-polarized). In both drawings, electric fields are shown as they would appear immediately adjacent to the conducting boundary. In (*a*) the field of a TE wave reverses direction upon reflection to establish a zero net field at the boundary. This constitutes a π phase shift, as is evident by considering a fictitious transmitted wave (*dashed line*), formed by a simple rotation of the reflected wave into alignment with the incident wave. In (*b*) an incident TM wave experiences a reversal of the *z component* of its electric field. The resultant field of the reflected wave, however, has not been phase-shifted; rotating the reflected wave into alignment with the incident wave (*dashed line*) shows this.

or 0, regardless of the angle of incidence. Thus the reflective phase shift has no bearing on the current problem, and we may simplify (3) to read:

$$\boxed{\kappa_m = \frac{m\pi}{d}} \tag{4}$$

which is valid for *both* TE and TM modes. Note from Fig. 14.9 that $\kappa_m = k\cos\theta_m$. Thus the wave angles for the allowed modes are readily found from (4) with (2):

$$\theta_m = \cos^{-1}\left(\frac{m\pi}{kd}\right) = \cos^{-1}\left(\frac{m\pi c}{\omega n d}\right) = \cos^{-1}\left(\frac{m\lambda}{2nd}\right) \tag{5}$$

where λ is the wavelength in free space.

We can next solve for the phase constant for each mode, using (1) with (4):

$$\beta_m = \sqrt{k^2 - \kappa_m^2} = k\sqrt{1 - \left(\frac{m\pi}{kd}\right)^2} = k\sqrt{1 - \left(\frac{m\pi c}{\omega n d}\right)^2} \tag{6}$$

We define the radian *cutoff frequency* for mode m as

$$\boxed{\omega_{cm} = \frac{m\pi c}{nd}} \tag{7}$$

so that (6) becomes

$$\boxed{\beta_m = \frac{n\omega}{c}\sqrt{1 - \left(\frac{\omega_{cm}}{\omega}\right)^2}} \tag{8}$$

The significance of the cutoff frequency is readily seen from (8): If the operating frequency, ω, is greater than the cutoff frequency for mode m, then that mode will have phase constant β_m that is real-valued, and so the mode will propagate. For $\omega < \omega_{cm}$, β_m is imaginary, and the mode does not propagate.

Associated with the cutoff frequency is the *cutoff wavelength*, λ_{cm}, defined as the free-space wavelength at which cutoff for mode m occurs. This will be

$$\boxed{\lambda_{cm} = \frac{2\pi c}{\omega_{cm}} = \frac{2nd}{m}} \tag{9}$$

Note, for example, that in an air-filled guide ($n = 1$) the wavelength at which the lowest-order mode first starts to propagate is $\lambda_{c1} = 2d$, or the plate separation is one-half wavelength. Mode m will propagate whenever $\omega > \omega_{cm}$, or equivalently whenever $\lambda < \lambda_{cm}$. Use of the cutoff wavelength enables us to construct a second useful form of Eq. (8):

$$\boxed{\beta_m = \frac{2\pi n}{\lambda}\sqrt{1 - \left(\frac{\lambda}{\lambda_{cm}}\right)^2}} \tag{10}$$

Example 14.1

A parallel-plate transmission line has plate separation $d = 1$ cm, and is filled with teflon, having dielectric constant $\epsilon'_R = 2.1$. Determine the maximum operating frequency such that only the TEM mode will propagate. Also find the range of frequencies over which the TE_1 and $TM_1 (m = 1)$ modes will propagate, and no higher order modes.

Solution. Using (7), the cutoff frequency for the first waveguide mode ($m = 1$), will be

$$f_{c1} = \frac{\omega_{c1}}{2\pi} = \frac{2.99 \times 10^{10}}{2\sqrt{2.1}} = 1.03 \times 10^{10} \text{ Hz} = 10.3 \text{ GHz}$$

To propagate only TEM waves, we must have $f < 10.3$ GHz. To allow TE_1 and TM_1 (along with TEM) only, the frequency range must be $\omega_{c1} < \omega < \omega_{c2}$, where $\omega_{c2} = 2\omega_{c1}$, from (7). Thus, the frequencies at which we will have the $m = 1$ modes and TEM will be $10.3\,\text{GHz} < f < 20.6\,\text{GHz}$.

Example 14.2

In the parallel-plate guide of Example 14.1, the operating wavelength is $\lambda = 2$ mm. How many waveguide modes will propagate?

Solution. For mode m to propagate, the requirement is $\lambda < \lambda_{cm}$. For the given waveguide and wavelength, the inequality becomes, using (9),

$$2 \text{ mm} < \frac{2\sqrt{2.1}\ (10\,\text{mm})}{m}$$

from which

$$m < \frac{2\sqrt{2.1}\ (10\,\text{mm})}{2\,\text{mm}} = 14.5$$

Thus the guide will support modes at the given wavelength up to order $m = 14$. Since there will be a TE and a TM mode for each value of m, this gives, in addition to the TEM mode, a total of 28 guided modes that are above cutoff.

The field configuration for a given mode can be found through the superposition of the fields of all the reflected waves. We can do this for the TE waves, for example, by writing the electric field phasor in the guide in terms of incident and reflected fields through

$$E_{ys} = E_0 e^{-j\mathbf{k_u}\cdot\mathbf{r}} - E_0 e^{-j\mathbf{k_d}\cdot\mathbf{r}} \tag{11}$$

where the wavevectors, $\mathbf{k_u}$ and $\mathbf{k_d}$, are indicated in Fig. 14.7. The minus sign in front of the second term arises from the π phase shift on reflection. From the geometry depicted in Fig. 14.9, we write

$$\mathbf{k_u} = \kappa_m \mathbf{a}_x + \beta_m \mathbf{a}_z \tag{12}$$

and

$$\mathbf{k_d} = -\kappa_m \mathbf{a}_x + \beta_m \mathbf{a}_z \tag{13}$$

Then, using

$$\mathbf{r} = x\mathbf{a}_x + z\mathbf{a}_z$$

Eq. (11) becomes

$$E_{ys} = E_0\left(e^{-j\kappa_m x} - e^{j\kappa_m x}\right)e^{-j\beta_m z} = 2jE_0 \sin(\kappa_m x)e^{-j\beta_m z} = E_0' \sin(\kappa_m x)e^{-j\beta_m z} \tag{14}$$

where the plane wave amplitude, E_0, and the overall phase are absorbed into E_0'. In real instantaneous form, (14) becomes

$$E_y(z,t) = \mathrm{Re}\left(E_{ys}e^{j\omega t}\right) = E_0' \sin(\kappa_m x)\cos(\omega t - \beta_m z) \qquad \text{(TE mode above cutoff)} \tag{15}$$

We interpret this as a wave that propagates in the positive z direction (down the guide) while having a field profile that varies with x.[3] The TE mode field is the *interference pattern* resulting from the superposition of the upward and downward plane waves. Note that if $\omega < \omega_{cm}$, then (8) yields an imaginary value for β_m, which we may write as $-j|\beta_m| = -j\alpha_m$. Eqs. (14) and (15) then become

$$E_{ys} = E_0' \sin(\kappa_m x)e^{-\alpha_m z} \tag{16}$$

$$E(z,t) = E_0' \sin(\kappa_m x)e^{-\alpha_m z}\cos(\omega t) \qquad \text{(TE mode below cutoff)} \tag{17}$$

This mode does not propagate, but simply oscillates at frequency ω, while exhibiting a field pattern that decreases in strength with increasing z. The attenuation coefficient, α_m, is found from (8) with $\omega < \omega_{cm}$:

$$\alpha_m = \frac{n\omega_{cm}}{c}\sqrt{1 - \left(\frac{\omega}{\omega_{cm}}\right)^2} = \frac{2\pi n}{\lambda_{cm}}\sqrt{1 - \left(\frac{\lambda_{cm}}{\lambda}\right)^2} \tag{18}$$

We note from (5) and (7) that the plane wave angle is related to the cutoff frequency and cutoff wavelength through

$$\boxed{\cos\theta_m = \frac{\omega_{cm}}{\omega} = \frac{\lambda}{\lambda_{cm}}} \tag{19}$$

So we see that at cutoff ($\omega = \omega_{cm}$), $\theta_m = 0$, and the plane waves are just reflecting back and forth over the cross section; they are making no forward progress down the guide. As ω is increased beyond cutoff (or λ is decreased), the wave angle increases, approaching 90° as ω approaches infinity (or as λ approaches zero). From Fig. 14.9, we have

$$\boxed{\beta_m = k\sin\theta_m = \frac{n\omega}{c}\sin\theta_m} \tag{20}$$

[3] We can also interpret this field as that of a standing wave in x, while it is a traveling wave in z.

and so the phase velocity of mode m will be

$$v_{pm} = \frac{\omega}{\beta_m} = \frac{c}{n \sin \theta_m} \tag{21}$$

The velocity minimizes at c/n for all modes, approaching this value at frequencies far above cutoff; v_{pm} approaches infinity as the frequency is reduced to approach the cutoff frequency. Again, phase velocity is the speed of the phases in the z direction, and the fact that this velocity may exceed the speed of light in the medium is not a violation of relativistic principles, as discussed in Sec. 12.5.

The energy will propagate at the group velocity, $v_g = d\omega/d\beta$. Using (8), we have

$$v_{gm}^{-1} = \frac{d\beta_m}{d\omega} = \frac{d}{d\omega}\left[\frac{n\omega}{c}\sqrt{1-\left(\frac{\omega_{cm}}{\omega}\right)^2}\,\right] \tag{22}$$

The derivative is straightforward. Carrying it out and taking the reciprocal of the result yields:

$$v_{gm} = \frac{c}{n}\sqrt{1-\left(\frac{\omega_{cm}}{\omega}\right)^2} = \frac{c}{n}\sin\theta_m \tag{23}$$

Group velocity is thus identified as the projection of the velocity associated with $\mathbf{k_u}$ or $\mathbf{k_d}$ into the z direction. This will be less than or equal to the velocity of light in the medium, c/n, as expected.

Example 14.3

In the guide of Example 14.1, the operating frequency is 25 GHz. Consequently, modes for which $m = 1$ and $m = 2$ will be above cutoff. Determine the *group delay difference* between these two modes over a distance of 1 cm. This is the difference in propagation times between the two modes when energy in each propagates over the 1 cm distance.

Solution. The group delay difference is expressed as

$$\Delta t = \left(\frac{1}{v_{g2}} - \frac{1}{v_{g1}}\right) \text{ (s/cm)}$$

From (23), along with the results of Example 14.1, we have

$$v_{g1} = \frac{c}{\sqrt{2.1}}\sqrt{1-\left(\frac{10.3}{25}\right)^2} = 0.63c$$

$$v_{g2} = \frac{c}{\sqrt{2.1}} \sqrt{1 - \left(\frac{20.6}{25}\right)^2} = 0.39c$$

Then

$$\Delta t = \frac{1}{c}\left[\frac{1}{.39} - \frac{1}{.63}\right] = 3.3 \times 10^{-11} \text{ s/cm} = 33 \text{ ps/cm}$$

This computation gives a rough measure of the *modal dispersion* in the guide, applying to the case of having only two modes propagating. A pulse, for example, whose center frequency is 25 GHz would have its energy divided between the two modes. The pulse would broaden by approximately 33 ps/cm of propagation distance as the energy in the modes separates. If, however, we include the TEM mode (as we really must), then the broadening will be even greater. The group velocity for TEM will be $c/\sqrt{2.1}$. The group delay difference of interest will then be between the TEM mode and the $m = 2$ mode (TE or TM). We would thus have

$$\Delta t_{net} = \frac{1}{c}\left[\frac{1}{.39} - 1\right] = 52 \text{ ps/cm}$$

✔ **D14.1.** Determine the wave angles, θ_m, for the first four modes ($m = 1,2,3,4$) in a parallel-plate guide with $d = 2$ cm, $\epsilon'_R = 1$, and $f = 30$ GHz.

Ans. 76°; 60°; 41°; 0°.

✔ **D14.2.** A parallel-plate guide has plate spacing $d = 5$ mm, and is filled with glass ($n = 1.45$). What is the maximum frequency at which the guide will operate in the TEM mode only?

Ans. 20.7 GHz.

✔ **D14.3.** A parallel-plate guide having $d = 1$ cm is filled with air. Find the cutoff wavelength for the $m = 2$ mode (TE or TM).

Ans. 1 cm.

14.3 PARALLEL-PLATE GUIDE ANALYSIS USING THE WAVE EQUATION

The most direct approach in the analysis of any waveguide is through the wave equation, which we solve subject to the boundary conditions at the conducting walls. The form of the equation that we will use is that of Eq. (12) in Sec. 11.1, which was written for the case of free-space propagation. We account for the dielectric properties in the waveguide by replacing k_0 in that equation with k to obtain:

$$\nabla^2 \mathbf{E}_s = -k^2 \mathbf{E}_s \qquad (24)$$

where $k = n\omega/c$ as before.

We can use the results of the last section to help visualize the process of solving the wave equation. For example, we may consider TE modes first, in which there will be only a y component of $\mathbf{E}$. The wave equation becomes:

$$\frac{\partial^2 E_{ys}}{\partial x^2} + \frac{\partial^2 E_{ys}}{\partial y^2} + \frac{\partial^2 E_{ys}}{\partial z^2} + k^2 E_{ys} = 0 \tag{25}$$

We assume that the width of the guide (in the y direction) is very large compared to the plate separation, d. Therefore we can assume no y variation in the fields (fringing fields are ignored), and so $\partial^2 E_{ys}/\partial y^2 = 0$. We also know that the z variation will be of the form $e^{-j\beta_m z}$. The form of the field solution will thus be

$$\boxed{E_{ys} = E_0 f_m(x) e^{-j\beta_m z}} \tag{26}$$

where E_0 is a constant, and where $f_m(x)$ is a normalized function to be determined (whose maximum value is unity). We have included subscript m on β, κ, and $f(x)$, since we anticipate several solutions that correspond to discrete modes, to which we associate mode number, m. We now substitute (26) into (25) to obtain:

$$\boxed{\frac{d^2 f_m(x)}{dx^2} + (k^2 - \beta_m^2) f_m(x) = 0} \tag{27}$$

where E_0 and $e^{-j\beta_m z}$ have divided out, and where we have used the fact that

$$\frac{d^2}{dz^2} e^{-j\beta_m z} = -\beta_m^2 e^{-j\beta_m z}$$

Note also that we have written (27) using the total derivative, d^2/dx^2 since f_m is a function only of x. We next make use of the geometry of Fig. 14.9, and note that $k^2 - \beta_m^2 = \kappa_m^2$. Using this in (27) we obtain

$$\boxed{\frac{d^2 f_m(x)}{dx^2} + \kappa_m^2 f_m(x) = 0} \tag{28}$$

The general solution of (28) will be

$$\boxed{f_m(x) = \cos(\kappa_m x) + \sin(\kappa_m x)} \tag{29}$$

We next apply the appropriate boundary conditions in our problem to evaluate κ_m. From Fig. 14.1, conducting boundaries appear at $x = 0$ and $x = d$, at which the tangential electric field (E_y) must be zero. In Eq. (29), only the $\sin(\kappa_m x)$ term will allow the boundary conditions to be satisfied, so

we retain it and drop the cosine term. The $x = 0$ condition is automatically satisfied by the sine function. The $x = d$ condition is met when we choose the value of κ_m such that

$$\kappa_m = \frac{m\pi}{d} \tag{30}$$

We recognize Eq. (30) as the same result that we obtained using the transverse resonance condition of the last section. The final form of E_{ys} is obtained by substituting $f_m(x)$ as expressed through (29) and (30) into (26), yielding a result that is consistent with the one expressed in Eq. (14):

$$\boxed{E_{ys} = E_0 \sin\left(\frac{m\pi x}{d}\right) e^{-j\beta_m z}} \tag{31}$$

An additional significance of the mode number m is seen when considering the form of the electric field of (31). Specifically, m is the number of spatial half-cycles of electric field that occur over the distance d in the transverse plane. This can be understood physically by considering the behavior of the guide at cutoff. As we learned in the last section, the plane wave angle of incidence in the guide at cutoff is zero, meaning that the wave simply bounces up and down between the conducting walls. The wave must be resonant in the structure, which means that the net round-trip phase shift is $2m\pi$. With the plane waves oriented vertically, $\beta_m = 0$, and so $\kappa_m = k = 2n\pi/\lambda_{cm}$. So at cutoff,

$$\frac{m\pi}{d} = \frac{2n\pi}{\lambda_{cm}} \tag{32}$$

which leads to

$$d = \frac{m\lambda_{cm}}{2n} \qquad \text{at cutoff} \tag{33}$$

Eq. (31) at cutoff then becomes

$$E_{ys} = E_0 \sin\left(\frac{m\pi x}{d}\right) = E_0 \sin\left(\frac{2n\pi x}{\lambda_{cm}}\right) \tag{34}$$

The waveguide is simply a one-dimensional *resonant cavity*, in which a wave can oscillate in the x direction if its wavelength as measured in the medium is an integer multiple of $2d$, where the integer is m.

Now, as the frequency increases, wavelength will decrease, and so the above requirement of wavelength equaling an integer multiple of $2d$ is no longer met. The response of the mode is to establish *z components* of $\mathbf{k_u}$ and $\mathbf{k_d}$ which results in the decreased wavelength being compensated by an increase in wavelength *as measured in the x direction*. Figure 14.12*a* and *b* show this effect for the $m = 4$ mode, in which the wave angle, θ_4, steadily increases with increasing frequency. Thus, the mode retains precisely the functional form of its field in the x direction, but establishes an increasing value of β_m as the frequency is raised. This invar-

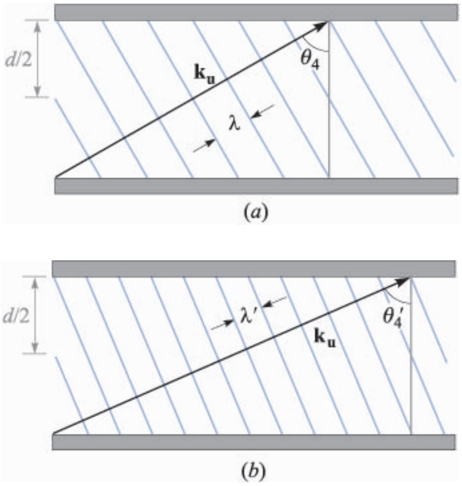


FIGURE 14.12
(*a*) A plane wave associated with an $m = 4$ mode, showing a net phase shift of 4π (two wavelengths measured in x) occurring over distance d in the transverse plane. (*b*) As frequency increases, an increase in wave angle is required to maintain the 4π transverse phase shift.

iance in the transverse spatial pattern means that the mode will retain its identity at all frequencies. Group velocity, expressed in (23), is changing as well, meaning that the changing wave angle with frequency is a mechanism for group velocity dispersion, known simply as *waveguide dispersion*. Pulses, for example, that propagate in a single waveguide mode will thus experience broadening in the manner considered in Sec. 12.6.

Having found the electric field, we can find the magnetic field using Maxwell's equations. We note from our plane wave model that we expect to obtain x and z components of $\mathbf{H_s}$ for a TE mode. We use the Maxwell equation

$$\nabla \times \mathbf{E_s} = -j\omega\mu\mathbf{H_s} \tag{35}$$

where, in the present case of having only a y component of $\mathbf{E_s}$, we have

$$\nabla \times \mathbf{E_s} = \frac{\partial E_{ys}}{\partial x}\mathbf{a}_z - \frac{\partial E_{ys}}{\partial z}\mathbf{a}_x = \kappa_m E_0 \cos(\kappa_m x)e^{-j\beta_m z}\mathbf{a}_z + j\beta_m E_0 \sin(\kappa_m x)e^{-j\beta_m z}\mathbf{a}_x \tag{36}$$

We solve for $\mathbf{H_s}$ by dividing both sides of (35) by $-j\omega\mu$. Performing this operation on (36), we obtain the two magnetic field components:

$$H_{xs} = -\frac{\beta_m}{\omega\mu}E_0 \sin(\kappa_m x)e^{-j\beta_m z} \tag{37}$$

$$H_{zs} = j\frac{\kappa_m}{\omega\mu}E_0 \cos(\kappa_m x)e^{-j\beta_m z} \tag{38}$$

Together, these two components form closed loop patterns for $\mathbf{H}_s$ in the x, z plane, as can be verified using the streamline plotting methods developed in Sec. 2.6.

It is interesting to consider the magnitude of $\mathbf{H}_s$, which is found through

$$|\mathbf{H}_s| = \sqrt{\mathbf{H}_s \cdot \mathbf{H}_s^*} = \sqrt{H_{xs}H_{xs}^* + H_{zs}H_{zs}^*} \tag{39}$$

Carrying this out using (37) and (38) results in

$$|\mathbf{H}_s| = \frac{E_o}{\omega\mu}\left(\kappa_m^2 + \beta_m^2\right)^{1/2}\left(\sin^2(\kappa_m x) + \cos^2(\kappa_m x)\right)^{1/2} \tag{40}$$

Using the fact that $\kappa_m^2 + \beta_m^2 = k^2$ and using the identity $\sin^2(\kappa_m x) + \cos^2(\kappa_m x) = 1$, (40) becomes

$$|\mathbf{H}_s| = \frac{k}{\omega\mu}E_0 = \frac{\omega\sqrt{\mu\epsilon}}{\omega\mu} = \frac{E_0}{\eta} \tag{41}$$

where $\eta = \sqrt{\mu/\epsilon}$. This result is consistent with our understanding of waveguide modes based on the superposition of plane waves, in which the relation between $\mathbf{E}_s$ and $\mathbf{H}_s$ is through the medium intrinsic impedance, η.

✔ **D14.4.** Determine the group velocity of the $m = 1$ (TE or TM) mode in an air-filled parallel-plate guide with $d = 0.5$ cm at $f =$ (a) 30 GHz, (b) 60 GHz, and (c) 100 GHz.

Ans. 0; 2.6×10^8 m/s; 2.9×10^8 m/s.

✔ **D14.5.** A TE mode in a parallel-plate guide is observed to have three maxima in its electric field pattern between $x = 0$ and $x = d$. What is the value of m?

Ans. 3.

14.4 RECTANGULAR WAVEGUIDES

In this section we consider the rectangular waveguide, a widely used structure that is usually used in the microwave region of the electromagnetic spectrum. A brief analysis of the structure will be presented here, with the goal of understanding the key operational features and special attributes of the guide. The reader is referred to Ref. 1 for further study.

The rectangular guide is shown in Fig. 14.2. We can relate this structure to that of the parallel-plate guide of the previous sections by thinking of it as two parallel-plate guides of orthogonal orientation that are assembled to form one unit. We thus have a pair of horizontal conducting walls (along the x direction) and a pair of vertical walls (along y), all of which now form one continuous boundary. The wave equation in its full three-dimensional form (Eq. (25)) must now be solved, since in general we may have field variations in all three coordinate directions. Assuming that the variation with z will be just $e^{-j\beta z}$ as before,

and assuming, for example, the existence of a y component of $\mathbf{E_s}$, Eq. (25) will take the form:

$$\frac{\partial^2 E_{ys}}{\partial x^2} + \frac{\partial^2 E_{ys}}{\partial y^2} + (k^2 - \beta_{mp}^2)E_{ys} = 0 \tag{42}$$

This equation brings to mind the two-dimensional Laplace equation problem of Sec. 7.5, in which the product solution method was used. The same basic method is used to solve (42) as well, and the resulting solution takes the general form:

$$E_{ys} = E_0 f_m(x) f_p(y) e^{-j\beta_{mp}z} \tag{43}$$

where f_m and f_p are sines or cosines. Two integers, m and p, are now needed to describe field variations in the x and y directions. Again, we are concerned with TE and TM modes, and the wave equation is solved separately for each type.

From here, the problem gets complicated and goes beyond the scope (and purpose) of our present treatment. Instead, much can be learned about this guide through our intuition, and through our knowledge of the parallel-plate guide. It turns out that the most important modes in the rectangular waveguide are of the same form as those of the parallel-plate structure. Consider, for example, the appearence of a TE mode in the rectangular guide. The electric field of such a mode could appear as shown in Fig. 14.13*a*, in which the field is vertically polarized, and terminates on the top and bottom plates. The field also becomes zero at the two vertical walls, as is required from our boundary condition on tangential electric field at a conducting surface. Let us consider the case in which the field exhibits no variation with y, but does vary with x and z (the latter according to $e^{-j\beta z}$). Consequently the $\partial^2/\partial y^2$ term in the wave equation (42) drops out and the equation becomes identical in form to (27), which was used for the parallel-plate guide. The field solutions are thus identical in form to (31), (37) and (38), but with a few minor notation differences:

$$E_{ys} = E_0 \sin(\kappa_{m0}x)e^{-j\beta_{m0}z} \tag{44}$$

$$H_{xs} = -\frac{\beta_{m0}}{\omega\mu} E_0 \sin(\kappa_{m0}x)e^{-j\beta_{m0}z} \tag{45}$$

$$H_{zs} = j\frac{\kappa_{m0}}{\omega\mu} E_0 \cos(\kappa_{m0}x)e^{-j\beta_{m0}z} \tag{46}$$

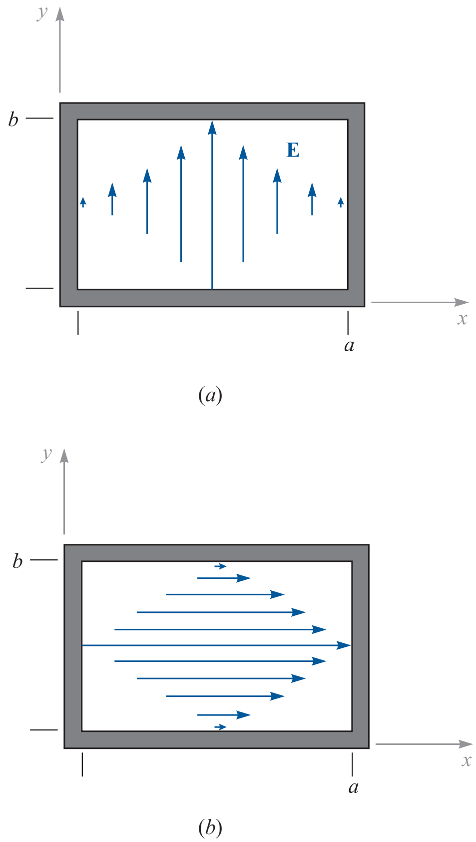


FIGURE 14.13
(*a*) TE_{10} and (*b*) TE_{01} mode electric field configurations in a rectangular waveguide.

where

$$\kappa_{m0} = \frac{m\pi}{a} \tag{47}$$

The fields of Eqs. (44) through (46) are those of a general mode of designation TE_{m0}, where the subscripts indicate that there are m half cycles of the electric field over the x dimension and zero variation in y. The phase constant is subscripted $m0$, and it is still true that

$$\kappa_{m0}^2 + \beta_{m0}^2 = k^2$$

The cutoff frequency for the TE_{m0} mode is given by (7), appropriately modified:

$$\boxed{\omega_c(m0) = \frac{m\pi c}{na}} \tag{48}$$

The phase constant, β_{m0}, is given by (8); all of the implications on mode behavior above and below cutoff are exactly the same as we found for the parallel-plate guide. The plane wave analysis is also carried out in the same manner. TE_{m0} modes can be modeled as plane waves that propagate down the guide by reflecting between the vertical side walls.

Another possibility is the TE_{0p} field configuration shown in Fig. 14.13*b*, which shows a horizontally polarized electric field. The wave equation (42) would now contain the $\partial^2/\partial y^2$ term and the $\partial^2/\partial x^2$ term would be dropped. The resulting fields would be those of (44) through (46) after a rotation through 90°, along with a few notational changes:

$$E_{xs} = E_0 \sin\left(\kappa_{0p} y\right) e^{-j\beta_{0p} z} \tag{49}$$

$$H_{ys} = \frac{\beta_{0p}}{\omega\mu} E_0 \sin\left(\kappa_{0p} y\right) e^{-j\beta_{0p} z} \tag{50}$$

$$H_{zs} = -j\frac{\kappa_{0p}}{\omega\mu} E_0 \cos(\kappa_{0p} y) e^{-j\beta_{0p} z} \tag{51}$$

where

$$\kappa_{0p} = \frac{p\pi}{b} \tag{52}$$

and where the cutoff frequency will be

$$\boxed{\omega_c(0p) = \frac{p\pi c}{nb}} \tag{53}$$

Other modes are possible which exhibit variations in both x and y. In general, the cutoff frequency for these modes is given by

$$\boxed{\omega_c(mp) = \sqrt{\left(\frac{m\pi c}{na}\right)^2 + \left(\frac{p\pi c}{nb}\right)^2}} \tag{54}$$

Modes having variations in both transverse directions include TE and TM, but only TE modes can have zero variation in x or y.

Of considerable practical interest is the mode that has the lowest cutoff frequency. If the guide dimensions are such that $b < a$, then inspection of (48) and (54) indicates that the lowest cutoff will occur for the TE_{10} mode. As such, this is the dominant (and most important) mode in the rectangular waveguide, since it can propagate alone if the operating frequency is appropriately chosen.

Example 14.4

An air-filled rectangular waveguide has dimensions $a = 2$ cm and $b = 1$ cm. Determine the range of frequencies over which the guide will operate single mode (TE_{10}).

Solution. Since the guide is air-filled, $n = 1$, and (48) gives, for $m = 1$:

$$f_c(10) = \frac{\omega_c}{2\pi} = \frac{c}{2a} = \frac{3 \times 10^{10}}{2(2)} = 7.5 \text{ GHz}$$

The next higher-order mode will be either TE_{20} or TE_{01}, which, from (48) and (54), will have the same cutoff frequency, since $a = 2b$. This frequency will be twice that found for TE_{10}, or 15 GHz. Thus the operating frequency range over which the guide will be single mode is 7.5 GHz $< f <$ 15 GHz.

Having seen how rectangular waveguides work, the questions arise: Why are they used and when are they useful? Let us consider for a moment the operation of a transmission line at frequencies high enough such that waveguide modes can occur. The onset of guided modes in a transmission line, known as *moding*, is in fact a problem that needs to be avoided, because signal distortion may result. A signal that is input to such a line will find its power divided in some proportions among the various modes. The signal power in each mode propagates at a group velocity unique to that mode. With the power thus distributed, distortion will occur over sufficient distances, as the signal components among the modes lose synchronization with each other, owing to the different delay times (group delays) associated with the different modes. We encountered this concept in Example 14.3.

The above problem of *modal dispersion* in transmission lines is avoided by ensuring that only the TEM mode propagates, and that all waveguide modes are below cutoff. This is accomplished either by using line dimensions that are smaller than one-half the signal wavelength, or by ensuring an upper limit to the operating frequency in a given line. But it is more complicated than this.

In Chap. 13, we saw that increasing the frequency increases the line loss as a result of the skin effect. This is manifested through the increase in the series resistance per unit length, R. One can compensate by increasing one or more dimensions in the line cross section, as Eqs. (17) and (33) in Chap. 13 show, but only to the point at which moding may occur. Typically, the increasing loss with increasing frequency will render the transmission line useless before the onset of moding, but one still cannot increase the line dimensions to reduce losses without considering the possibility of moding. This limitation on dimensions also limits the power-handling capability of the line, since the voltage at which dielectric breakdown occurs decreases with decreasing conductor separation. Consequently, the use of transmission lines as frequencies increase beyond a certain point becomes undesirable, since losses will become excessive, and since the limitation on dimensions will limit the power-handling capability. Instead, we look to other guiding structures, among which is the rectangular guide.

The important fundamental difference between the rectangular waveguide (or any hollow pipe guide) and the transmission line is that the rectangular guide *will not support a TEM mode*. We have already demonstrated this in our study of the TE wave. The fact that the guide is formed from a completely enclosed metal structure means that any electric field distribution in the transverse plane must exhibit variations in the plane; this is because all electric field components that are tangent to the conductors must be zero at the conducting boundaries. Since **E** varies in the transverse plane, the computation of **H** through $\nabla \times \mathbf{E} = -j\omega\mu\mathbf{H}$ *must* lead to a z component of **H**, and so we cannot have a TEM mode. We cannot find any other orientation of a completely transverse **E** in the guide that will allow a completely transverse **H**.

Since the rectangular guide will not support a TEM mode, it will not operate until the frequency exceeds the cutoff frequency of the lowest-order guided mode of the structure. Thus, it must be constructed of large enough size to accomplish this for a given frequency; the required transverse dimensions will consequently be larger than those of a transmission line that is designed to support only the TEM mode. The increased size, coupled with the fact that there is more conductor surface area than in a transmission line of equal volume means that losses will be substantially lower in the rectangular waveguide structure. Additionally, the guides will support more power at a given electric field strength than a transmission line, since the rectangular guide will have a higher cross-sectional area.

Still, hollow pipe guides must operate in a single mode in order to avoid the signal distortion problems arising from multimode transmission. This means that the guides must be of dimension such that they operate above the cutoff frequency of the lowest-order mode, but below the cutoff frequency of the next higher-order mode, as demonstrated in Example 14.4. Increasing the operating frequency again means that the guide transverse dimensions must be decreased to maintain single-mode operation. This can be accomplished to a point at which skin effect losses again become problematic (remember that the skin depth is decreasing with increasing frequency, in addition to the decrease in metal surface area with diminishing guide size). In addition, the guides become too difficult to fabricate, with machining tolerances becoming more stringent. So again, as frequencies are further increased, we look for another type of structure.

✔ **D14.6.** Specify the minimum width, a, and the maximum height, b, of an air-filled rectangular guide so that it will operate single mode over the frequency range $15\,\text{GHz} < f < 20\,\text{GHz}$.

Ans. 1 cm; 0.75 cm.

14.5 DIELECTRIC WAVEGUIDES

When skin effect losses become excessive, a good way to remove them is to remove the metal in the structure entirely, and use interfaces between dielectrics

for the confining surfaces. We thus obtain a *dielectric waveguide*; a basic form, the *symmetric slab waveguide*, is shown in Fig. 14.14. The structure is so-named because of its vertical symmetry about the z axis. The guide is assumed to have width in y much greater than the slab thickness, d, so the problem becomes two-dimensional, with fields presumed to vary with x and z while being independent of y. The slab guide works in very much the same way as the parallel plate waveguide, except wave reflections occur at the interfaces between dielectrics, having different refractive indices, n_1 for the slab, and n_2 for the surrounding regions above and below. In the dielectric guide, total reflection is needed, so the incident angle must exceed the critical angle. Consequently, as discussed in Sec. 12.5, the slab index, n_1, must be greater than that of the surrounding materials, n_2. Dielectric guides differ from conducting guides in that power is not completely confined to the slab, but resides partially above and below.

Dielectric guides are used primarily at optical frequencies (on the order of 10^{14} Hz). Again, guide transverse dimensions must be kept on the order of a wavelength to achieve operation in a single mode. A number of fabrication methods can be used to accomplish this. For example, a glass plate can be doped with materials that will raise the refractive index. The doping process allows materials to be introduced only within a thin layer adjacent to the surface that is a few micrometers thick.

To understand how the guide operates, consider Fig. 14.15, which shows a wave propagating through the slab by multiple reflections, but where *partial transmission* into the upper and lower regions occurs at each bounce. Wavevectors are shown in the middle and upper regions, along with their components in the x and z directions. As we found in Chap. 12, the z components (β) of all wavevectors are equal, as must be true if the field boundary conditions at the interfaces are to be satisfied for all positions and times. Partial transmission at the boundaries is, of course, an undesirable situation, since power in the slab will eventually leak away. We thus have a *leaky wave* propagating in the struc-

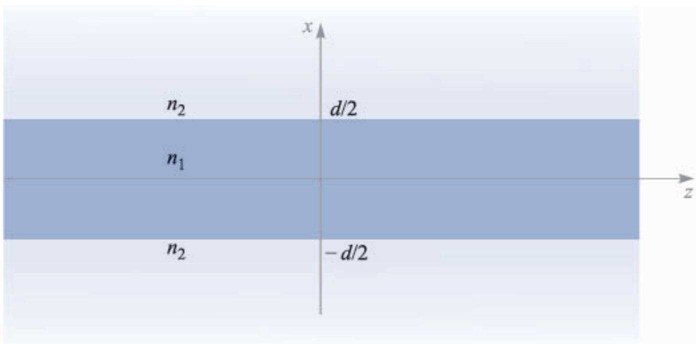


FIGURE 14.14
Symmetric dielectric slab waveguide structure, where waves propagate along z. The guide is assumed infinite in the y direction, thus making the problem two-dimensional.

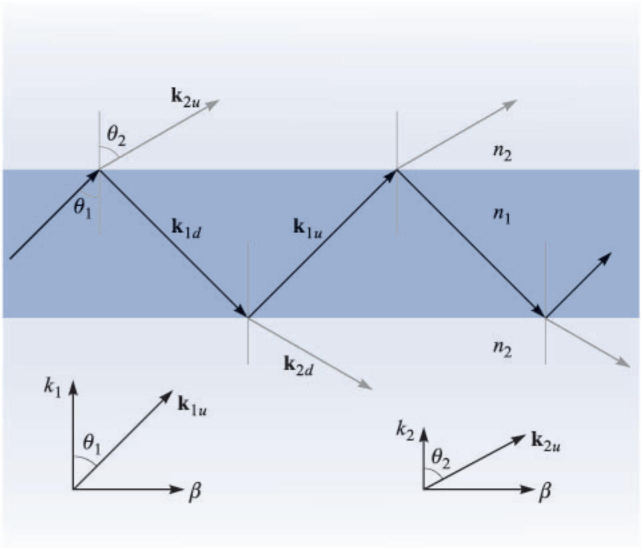


FIGURE 14.15
Plane wave geometry of a leaky wave in a symmetric slab waveguide. For a guided mode, total reflection occurs in the interior, and the x components of $\mathbf{k}_{2u}$ and $\mathbf{k}_{2d}$ are imaginary.

ture, whereas we need to have a guided mode. Note that in either case, we still have the two possibilities on wave polarization, and the resulting mode designation—either TE or TM.

Total power reflection at the boundaries for TE or TM waves implies, respectively, that $|\Gamma_s|^2$ or $|\Gamma_p|^2$ is unity, where the reflection coefficients are given in Eqs. (71) and (69) in Chap. 12:

$$\Gamma_s = \frac{\eta_{2s} - \eta_{1s}}{\eta_{2s} + \eta_{1s}} \tag{55}$$

and

$$\Gamma_p = \frac{\eta_{2p} - \eta_{1p}}{\eta_{2p} + \eta_{1p}} \tag{56}$$

As discussed in Sec. 12.5, we require that the effective impedances, η_{2s} or η_{2p} be purely imaginary, zero, or infinite if (55) or (56) are to have unity magnitudes. Knowing that

$$\eta_{2s} = \frac{\eta_2}{\cos\theta_2} \tag{57}$$

and

$$\eta_{2p} = \eta_2 \cos\theta_2 \tag{58}$$

the requirement is that $\cos\theta_2$ be zero or imaginary, where, from Eq. (75), Sec. 12.5,

$$\cos\theta_2 = \left[1 - \sin^2\theta_2\right]^{1/2} = \left[1 - \left(\frac{n_1}{n_2}\right)^2 \sin^2\theta_1\right]^{1/2} \tag{59}$$

As a result, we require that

$$\theta_1 \geq \theta_c \tag{60}$$

where the critical angle is defined through

$$\sin\theta_c = \frac{n_2}{n_1} \tag{61}$$

Now, from the geometry of Fig. 14.15, we can construct the field distribution of a TE wave in the guide using plane wave superposition. In the slab region ($-d/2 < x < d/2$), we have

$$E_{y1s} = E_0 e^{-j\mathbf{k}_{1u}\cdot\mathbf{r}} \pm E_0 e^{-j\mathbf{k}_{1d}\cdot\mathbf{r}} \qquad \left(-\frac{d}{2} < x < \frac{d}{2}\right) \tag{62}$$

where

$$\mathbf{k}_{1u} = \kappa_1 \mathbf{a_x} + \beta \mathbf{a_z} \tag{63}$$

and

$$\mathbf{k}_{1d} = -\kappa_1 \mathbf{a_x} + \beta \mathbf{a_z} \tag{64}$$

The second term in (62) may either add to or subtract from the first term, since either case would result in a symmetric intensity distribution in the x direction. We expect this, since the guide is symmetric. Now, using $\mathbf{r} = x\mathbf{a_x} + z\mathbf{a_z}$, (62) becomes

$$\boxed{E_{y1s} = E_0[e^{j\kappa_1 x} + e^{-j\kappa_1 x}]e^{-j\beta z} = 2E_0\cos(\kappa_1 x)e^{-j\beta z}} \tag{65}$$

for the choice of the plus sign in (62), and

$$\boxed{E_{y1s} = E_0[e^{j\kappa_1 x} - e^{-j\kappa_1 x}]e^{-j\beta z} = 2jE_0\sin(\kappa_1 x)e^{-j\beta z}} \tag{66}$$

if the minus sign is chosen. Since $\kappa_1 = n_1 k_0 \cos\theta_1$, we see that larger values of κ_1 imply smaller values of θ_1 at a given frequency. In addition, larger κ_1 values result in a greater number of spatial oscillations of the electric field over the transverse dimension, as (65) and (66) show. We found similar behavior in the parallel-plate guide. In the slab waveguide, as with the parallel-plate guide, we associate higher-order modes with increasing values of κ_1.[4]

[4] It would be appropriate to add the mode number subscript, m, to κ_1, κ_2, β, and θ_1, since, as was true with the metal guides, we will obtain discrete values of these quantities. To keep notation simple, the m subscript is suppressed, and we will assume it to be understood. Again, the subscripts 1 and 2 in this section indicate respectively the slab and surrounding *regions*, and have nothing to do with mode number.

In the regions above and below the slab, waves propagate according to wavevectors $\mathbf{k}_{2u}$ and $\mathbf{k}_{2d}$ as shown in Fig. 14.15. For example, above the slab ($x > d/2$), the TE electric field will be of the form:

$$E_{y2s} = E_{02}e^{-j\mathbf{k}_2\cdot\mathbf{r}} = E_{02}e^{-j\kappa_2 x}e^{-j\beta z} \tag{67}$$

However, $\kappa_2 = n_2 k_0 \cos\theta_2$, where $\cos\theta_2$, given in (59), is imaginary. We may therefore write

$$\kappa_2 = -j\gamma_2 \tag{68}$$

where γ_2 is real, and is given by (using 59)

$$\gamma_2 = j\kappa_2 = jn_2k_0\cos\theta_2 = jn_2k_0(-j)\left[\left(\frac{n_1}{n_2}\right)^2\sin^2\theta_1 - 1\right]^{1/2} \tag{69}$$

Eq. (67) now becomes

$$\boxed{E_{y2s} = E_{02}e^{-\gamma_2(x-d/2)}e^{-j\beta z} \qquad \left(x > \frac{d}{2}\right)} \tag{70}$$

where the x variable in (67) has been replaced by $x - (d/2)$ to position the field magnitude, E_{02}, at the boundary. Using similar reasoning, the field in the region below the lower interface, where x is negative, and where $\mathbf{k}_{2d}$ is involved, will be

$$\boxed{E_{y2s} = E_{02}e^{\gamma_2(x+d/2)}e^{-j\beta z} \qquad \left(x < -\frac{d}{2}\right)} \tag{71}$$

The fields expressed in (70) and (71) are those of *surface* waves. Note that they propagate in the z direction only, according to $e^{-j\beta z}$, but simply reduce in amplitude with increasing $|x|$, according to the $e^{-\gamma_2(x-d/2)}$ term in (70) and the $e^{\gamma_2(x+d/2)}$ term in (71). These waves represent a certain fraction of the total power in the mode, and so we see an important fundamental difference between dielectric waveguides and metal waveguides: in the dielectric guide, the fields (and guided power) exist over a cross section that extends beyond the confining boundaries, and in principle exist over an infinite cross section. In practical situations, the exponential decay of the fields above and below the boundaries is typically sufficient to render the fields negligible within a few slab thicknesses from each boundary.

The total electric field distribution is composed of the field in all three regions, and is sketched in Fig. 14.16 for the first few modes. Within the slab, the field is oscillatory and is of a similar form to that of the parallel-plate waveguide. The difference is that the fields in the slab guide do not reach zero at the boundaries, but connect to the evanescent fields above and below the slab. The restriction is that the TE fields on either side of a boundary (being tangent to the interface) must match at the boundary. Specifically,

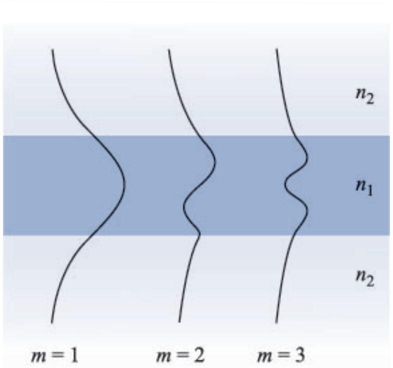


FIGURE 14.16
Electric field amplitude distributions over the transverse plane for the first three modes in a symmetric slab guide.

$$E_{y1s}|_{x=\pm d/2} = E_{y2s}|_{x=\pm d/2} \tag{72}$$

Applying the above condition to (65), (66), (70), and (71) results in the final expressions for the TE electric field in the symmetric slab waveguide, for the cases of even and odd symmetry:

$$E_{se}(\text{even TE}) = \begin{cases} E_{0e}\cos(\kappa_1 x)e^{-j\beta z} & \left(-\frac{d}{2} < x < \frac{d}{2}\right) \\ E_{0e}\cos\left(\kappa_1 \frac{d}{2}\right)e^{-\gamma_2(x-d/2)}e^{-j\beta z} & \left(x > \frac{d}{2}\right) \\ E_{0e}\cos\left(\kappa_1 \frac{d}{2}\right)e^{\gamma_2(x+d/2)}e^{-j\beta z} & \left(x < -\frac{d}{2}\right) \end{cases} \tag{73}$$

$$E_{so}(\text{odd TE}) = \begin{cases} E_{0o}\sin(\kappa_1 x)e^{-j\beta z} & \left(-\frac{d}{2} < x < \frac{d}{2}\right) \\ E_{0o}\sin\left(\kappa_1 \frac{d}{2}\right)e^{-\gamma_2(x-d/2)}e^{-j\beta z} & \left(x > \frac{d}{2}\right) \\ -E_{0o}\sin\left(\kappa_1 \frac{d}{2}\right)e^{\gamma_2(x+d/2)}e^{-j\beta z} & \left(x < -\frac{d}{2}\right) \end{cases} \tag{74}$$

Solution of the wave equation yields (as it must) results identical to these. The reader is referred to Refs. 2 and 3 for the details. The magnetic field for the TE modes will consist of x and z components, as was true for the parallel-plate guide. Finally, the TM mode fields will be nearly the same in form as those of TE modes, but with a simple rotation in polarization of the plane wave components by 90°. Thus, in TM modes, H_y will result, and will have the same form as E_y for TE, as presented in (73) and (74).

Apart from the differences in the field structures, the dielectric slab waveguide operates in a manner that is qualitatively similar to the parallel-plate guide. Again, a finite number of discrete modes will be allowed at a given frequency, and this number increases as frequency increases. Higher order modes are characterized by successively smaller values of θ_1.

An important difference in the slab guide occurs at cutoff for any mode. We know that $\theta = 0$ at cutoff in the metal guides. In the dielectric guide at cutoff, the wave angle, θ_1, is equal to the *critical angle*, θ_c. Then, as the frequency of a given mode is raised, its θ_1 value increases beyond θ_c in order to maintain transverse resonance, while maintaining the same number of field oscillations in the transverse plane.

As wave angle increases, however, the character of the evanescent fields changes significantly. This can be understood by considering the wave angle dependence on evanescent decay coefficient, γ_2, as given by (69). Note in that equation, that as θ_1 increases (as frequency goes up), γ_2 also increases, leading to a more rapid fall-off of the fields with increasing distance above and below the slab. Hence, the mode becomes more tightly confined to the slab as frequency is raised. Also, at a given frequency, higher-order modes, having smaller wave angles, will have lower values of γ_2 as (69) indicates. Consequently, when considering several modes propagating together at a single frequency, the higher order modes will carry a greater percentage of their power in the upper and lower regions surrounding the slab, than modes of lower order.

One can determine the conditions under which modes will propagate by using the transverse resonance condition, as we did with the parallel-plate guide. We perform the transverse round-trip analysis in the slab region in the same manner as was done in Sec. 14.2, and obtain an equation similar to (3):

$$\kappa_1 d + \phi_{TE} + \kappa_1 d + \phi_{TE} = 2m\pi \tag{75}$$

for TE waves and

$$\kappa_1 d + \phi_{TM} + \kappa_1 d + \phi_{TM} = 2m\pi \tag{76}$$

for the TM case. The phase shifts on reflection, ϕ_{TE} and ϕ_{TM} are the phases of the reflection coefficients, Γ_s and Γ_p, given in (55) and (56). These are readily found, but turn out to be functions of θ_1. As we know, κ_1 also depends on θ_1, but in a different way than ϕ_{TE} and ϕ_{TM}. Consequently, (75) and (76) are *transcendental* in θ_1, and cannot be solved in closed form. Instead, numerical or graphical methods must be used (see Refs. 2 or 3). Emerging from this solution process, however, is a fairly simple cutoff condition for any TE or TM mode:

$$k_0 d\sqrt{n_1^2 - n_2^2} \geq (m-1)\pi \qquad (m = 1, 2, 3, \ldots) \tag{77}$$

For mode m to propagate, (77) must hold. The physical interpretation of the mode number m is again the number of half-cycles of the electric field (for TE modes) or magnetic field (for TM modes) that occur over the transverse dimension. The lowest order mode ($m = 1$) is seen to have no cutoff—it will propagate from zero frequency on up. We will thus achieve single mode operation (actually a single pair of TE and TM modes) if we can ensure that the $m = 2$ modes are below cutoff. Using (77), our single mode condition will thus be:

$$k_0 d\sqrt{n_1^2 - n_2^2} < \pi \tag{78}$$

Using $k_0 = 2\pi/\lambda$, the wavelength range over which single mode operation occurs is

$$\lambda > 2d\sqrt{n_1^2 - n_2^2} \tag{79}$$

Example 14.5

A symmetric dielectric slab waveguide is to guide light at wavelength $\lambda = 1.30\ \mu$m. The slab thickness is to be $d = 5.00\ \mu$m, and the refractive index of the surrounding material is $n_2 = 1.450$. Determine the maximum allowable refractive index of the slab material that will allow single TE and TM mode operation.

Solution. Eq. (79) can be rewritten in the form:

$$n_1 < \sqrt{\left(\frac{\lambda}{2d}\right)^2 + n_2^2}$$

So

$$n_1 < \sqrt{\left(\frac{1.30}{2(5.00)}\right)^2 + (1.450)^2} = 1.456$$

Clearly, fabrication tolerances are very exacting when constructing dielectric guides for single mode operation!

Optical fiber waveguides work on the same principle as the basic dielectric guide, except, of course for the round cross section. A *step index* fiber is shown in Fig. 14.5 in which a high index *core* of radius a is surrounded by a lower-index *cladding* of radius b. Light is confined to the core through the mechanism of total reflection, but again some fraction of the power resides in the cladding as well. As we found in the slab guide, the cladding power again moves in toward the core as frequency is raised. Additionally, as is true in the slab waveguide, the fiber supports a mode that has no cutoff. Again, without proof (but see Refs. 2 or 3), the condition for single-mode operation in a step index fiber is found to be similar to that in the slab:

$$\lambda > \frac{2\pi a}{2.405}\sqrt{n_1^2 - n_2^2} \tag{80}$$

Typical dimensions of a single mode fiber include core diameters between 5 and 10 μm, with the cladding diameter usually 125 μm. As we found in the above example, refractive index differences between core and cladding are quite small, and are typically a small fraction of a percent.

✔ **D14.7.** A 0.5 mm thick slab of glass ($n_1 = 1.45$) is surrounded by air ($n_2 = 1$). The slab guides infrared light at wavelength $\lambda = 1.0\ \mu$m. How many TE and TM modes will propagate?

Ans. 2102.

14.6 BASIC ANTENNA PRINCIPLES

In this final section we explore some concepts on radiation of electromagnetic energy from a simple dipole antenna. A complete discussion of antennas and their applications would require several chapters or entire books. Our purpose is to produce a fundamental understanding of how electromagnetic fields radiate from *current distributions*. So for the first time, we shall have the specific field which results from a specific time-varying *source*. In the discussion of waves and fields in bulk media and in waveguides, only the wave motion in the medium was investigated, and the sources of the fields were not considered. The current distribution in a conductor was a similar problem, although we did at least relate the current to an assumed electric field intensity at the conductor surface. This might be considered as a source, but it is not a very practical one for it is infinite in extent.

We now assume a current filament (of infinitesimally-small cross section) as the source, positioned within an infinite lossless medium. The filament is taken as a differential length, but we shall be able to extend the results easily to a filament which is short compared to a wavelength, specifically less than about one-quarter of a wavelength overall. The differential filament is shown at the origin and is oriented along the z axis in Fig. 14.17. The positive sense of the current is taken in the $\mathbf{a}_z$ direction. We assume a uniform current $I_0 \cos \omega t$ in this short length d and do not concern ourselves at present with the apparent discontinuity at each end.

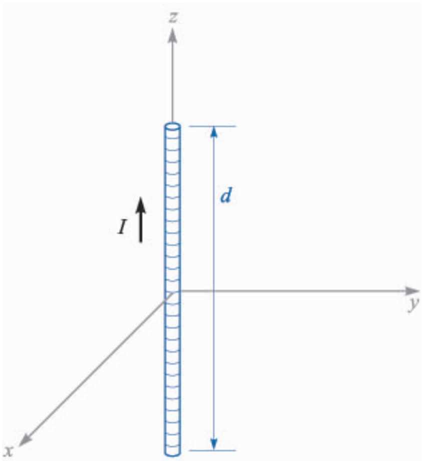


FIGURE 14.17
A differential current filament of length d carries a current $I = I_0 \cos \omega t$.

We shall not attempt at this time to discover the "source of the source," but shall merely assume that the current distribution cannot be changed by any field which it produces.

The first step is the application of the retarded vector magnetic potential expression, as presented in Sec. 10.5,

$$\mathbf{A} = \int \frac{\mu[I]d\mathbf{L}}{4\pi R}$$

where $[I]$ is a function of the retarded time $t - R/v$. When a single frequency is used to drive the antenna, v is the phase velocity at that frequency. Since no integration is required for the very short filament assumed, we have

$$\mathbf{A} = \frac{\mu[I]d}{4\pi R}\mathbf{a}_z$$

Only the z component of $\mathbf{A}$ is present, for the current is only in the $\mathbf{a}_z$ direction. At any point P, distant R from the origin, the vector potential is retarded by R/v and

$$I = I_0 \cos \omega t$$

becomes

$$[I] = I_0 \cos\left[\omega\left(t - \frac{R}{v}\right)\right]$$

$$[I_s] = I_0 e^{-j\omega R/v}$$

Thus

$$A_{zs} = \frac{\mu I_0 d}{4\pi R} e^{-j\omega R/v}$$

Using a mixed coordinate system for the moment, let us replace R by the small r of the spherical coordinate system and then determine which spherical components are represented by A_{zs}. Figure 14.18 helps to determine that

$$A_{rs} = A_{zs} \cos\theta$$

$$A_{\theta s} = -A_{zs} \sin\theta$$

and therefore

$$A_{rs} = \frac{\mu I_0 d}{4\pi r} \cos\theta \, e^{-j\omega r/v}$$

$$A_{\theta s} = -\frac{\mu I_0 d}{4\pi r} \sin\theta \, e^{-j\omega r/v}$$

From these two components of the vector magnetic potential at P we may find $\mathbf{B}_s$ or $\mathbf{H}_s$ from the definition of $\mathbf{A}_s$,

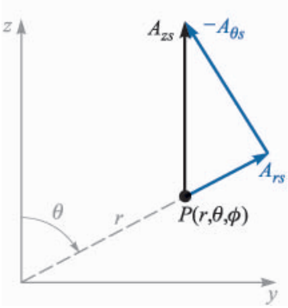


FIGURE 14.18
The resolution of A_{zs} at $P(r, \theta, \phi)$ into the two spherical components A_{rs} and $A_{\theta s}$. The sketch is arbitrarily drawn in the $\phi = 90^\circ$ plane.

$$\mathbf{B_s} = \mu \mathbf{H_s} = \nabla \times \mathbf{A_s}$$

by merely taking the indicated partial derivatives. Thus

$$H_{\phi s} = \frac{1}{\mu r}\frac{\partial}{\partial r}(rA_{\theta s}) - \frac{1}{\mu r}\frac{\partial A_{rs}}{\partial \theta}$$

$$H_{rs} = H_{\theta s} = 0$$

and

$$H_{\phi s} = \frac{I_0 d}{4\pi}\sin\theta\, e^{-j\omega r/v}\left(j\frac{\omega}{vr} + \frac{1}{r^2}\right)$$

The components of the electric field which must be associated with this magnetic field are found from the point form of Ampere's circuital law as it applies to a region in which conduction and convection current are absent,

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}$$

or in complex notation,

$$\nabla \times \mathbf{H_s} = j\omega\epsilon \mathbf{E_s}$$

Expansion of the curl in spherical coordinates leads to

$$E_{rs} = \frac{1}{j\omega\epsilon}\frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(H_{\phi s}\sin\theta)$$

$$E_{\theta s} = \frac{1}{j\omega\epsilon}\left(-\frac{1}{r}\right)\frac{\partial}{\partial r}(rH_{\phi s})$$

or

$$E_{rs} = \frac{I_0 d}{2\pi}\cos\theta\, e^{-j\omega r/v}\left(\frac{1}{\epsilon v r^2} + \frac{1}{j\omega\epsilon r^3}\right)$$

$$E_{\theta s} = \frac{I_0 d}{4\pi}\sin\theta\, e^{-j\omega r/v}\left(\frac{j\omega}{\epsilon v^2 r} + \frac{1}{\epsilon v r^2} + \frac{1}{j\omega\epsilon r^3}\right)$$

In order to simplify the interpretation of the terms enclosed in parentheses above, we make the substitutions $\omega = 2\pi f$, $f\lambda = v$, $v = 1/\sqrt{\mu\epsilon}$, and $\eta = \sqrt{\mu/\epsilon}$, producing

$$\boxed{H_{\phi s} = \frac{I_0 d}{4\pi} \sin\theta \, e^{-j2\pi r/\lambda} \left(j\frac{2\pi}{\lambda r} + \frac{1}{r^2} \right)} \tag{81}$$

$$\boxed{E_{rs} = \frac{I_0 d \eta}{2\pi} \cos\theta \, e^{-j2\pi r/\lambda} \left(\frac{1}{r^2} + \frac{\lambda}{j2\pi r^3} \right)} \tag{82}$$

$$\boxed{E_{\theta s} = \frac{I_0 d \eta}{4\pi} \sin\theta \, e^{-j2\pi r/\lambda} \left(j\frac{2\pi}{\lambda r} + \frac{1}{r^2} + \frac{\lambda}{j2\pi r^3} \right)} \tag{83}$$

These three equations are indicative of the reason that so many problems involving antennas are solved by experimental rather than theoretical methods. They have resulted from three general steps: an integration (atypically trivial) and two differentiations. These steps are sufficient to cause the simple current element and its simple current expression to "blow up" into the complicated field described by (81) to (83). In spite of this complexity, several interesting observations are possible.

We might notice first the $e^{-j2\pi r/\lambda}$ factor appearing with each component. This indicates propagation outward from the origin in the positive r direction with a phase factor $\beta = 2\pi/\lambda$; thus the wavelength is λ and the velocity $v = 1/\sqrt{\mu\epsilon}$. We use the term "wavelength" now in a somewhat broader sense than the original definition, which identified the wavelength of a uniform plane wave with the distance between two points, measured in the direction of propagation, at which the wave has identical instantaneous values. Here there are additional complications caused by the terms enclosed in parentheses, which are complex functions of r. These variations must now be neglected in determining the wavelength. This is equivalent to a determination of the wavelength at a large distance from the origin, and we may demonstrate this by sketching the H_ϕ component as a function of r under the following conditions:

$$I_0 d = 4\pi \qquad \theta = 90^\circ \qquad t = 0 \qquad f = 300 \text{ MHz}$$

$$v = 3 \times 10^8 \text{ m/s (free space)} \qquad \lambda = 1 \text{ m}$$

Therefore

$$H_{\phi s} = \left(j\frac{2\pi}{r} + \frac{1}{r^2} \right) e^{-j2\pi r}$$

and the real part may be determined at $t = 0$,

$$H_\phi = \sqrt{\left(\frac{2\pi}{r}\right)^2 + \frac{1}{r^4}}\, \cos[(\tan^{-1} 2\pi r) - 2\pi r]$$

Knowing that $\cos(a - b) = \cos a \cos b + \sin a \sin b$ and that $\cos(\tan^{-1} x) = 1/\sqrt{1 + x^2}$, we may simplify this result to

$$H_\phi = \frac{1}{r^2}(\cos 2\pi r + 2\pi r \sin 2\pi r)$$

Values obtained from this last equation are plotted against r in the range $1 \le r \le 2$ in Fig. 14.19*a*; the curve is noticeably nonsinusoidal. At $r = 1$, $H_\phi = 1$, while at $r = 2$, one wavelength greater, $H_\phi = 0.25$. Moreover, the curve crosses the axis (with positive slope) at $r = 1 - 0.0258$ and $r = 2 - 0.0127$, again a distance not equal to a wavelength. If a similar sketch is made in the range $101 \le r \le 102$, shown in Fig. 14.19*b* on a different amplitude scale, an essentially sinusoidal wave is obtained and the instantaneous values of H_ϕ at $r = 101$ and $r = 102$ are 0.000 099 8 and 0.000 099 6. The maximum amplitudes of the positive and negative portions of the waveform differ by less than 1 percent, and we may say that for all practical purposes the wave in this region is a uniform plane wave having a sinusoidal variation with distance (and time, of course) and a well-defined wavelength. This wave evidently carries energy away from the differential antenna, and we shall calculate this power shortly.

Continuing the investigation of (81) to (83), let us now take a more careful look at the expressions containing terms varying as $1/r^3$, $1/r^2$, and $1/r$. At points very close to the current element the $1/r^3$ term must be dominant. In the numerical example we have used, the relative values of the terms in $1/r^3$, $1/r^2$, and $1/r$

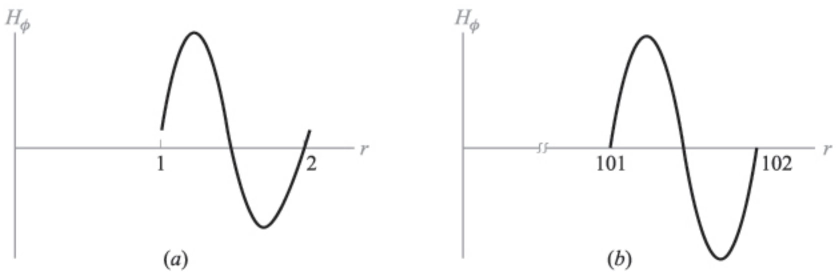


FIGURE 14.19
The instantaneous amplitude of H_ϕ for the special case of a current element having $I_0 d = 4\pi$ and $\lambda = 1$ is plotted at $\theta = 90^\circ$ and $t = 0$ (*a*) in the region $1 \le r \le 2$ near the antenna, and (*b*) in the region $101 \le r \le 102$ distant from the antenna. The left curve is noticeably nonsinusoidal, since if it were a sinusoid, its end points should reach the r axis precisely at 1 and 2.

in the $E_{\theta s}$ expression are about 250, 16, and 1, respectively, when r is 1 cm. The variation of an electric field as $1/r^3$ should remind us of the *electrostatic* field of the dipole (Chap. 4). This term represents energy stored in a reactive (capacitive) field, and it does not contribute to the radiated power. The inverse-square term in the $H_{\phi s}$ expression is similarly important only in the region very near to the current element and corresponds to the *induction* field of the dc element given by the Biot-Savart law. At distances corresponding to 10 or more wavelengths from the oscillating current element, all terms except the inverse-distance $(1/r)$ term may be neglected and the *distant* or *radiation* fields become

$$\boxed{E_{rs} = 0}$$

$$\boxed{E_{\theta s} = j\frac{I_0 d\eta}{2\lambda r}\sin\theta\, e^{-j2\pi r/\lambda}} \tag{84}$$

$$\boxed{H_{\phi s} = j\frac{I_0 d}{2\lambda r}\sin\theta\, e^{-j2\pi r/\lambda}} \tag{85}$$

or

$$\boxed{E_{\theta s} = \eta H_{\phi s}}$$

The relationship between $E_{\theta s}$ and $H_{\phi s}$ is thus seen to be that between the electric and magnetic fields of the uniform plane wave, thus substantiating the conclusion we reached when investigating the wavelength.

The variation of both radiation fields with the polar angle θ is the same; the fields are maximum in the equatorial plane (xy plane) of the current element and vanish off the ends of the element. The variation with angle may be shown by plotting a *vertical pattern* (assuming a vertical orientation of the current element) in which the relative magnitude of $E_{\theta s}$ is plotted against θ for a constant r. The pattern is usually shown on polar coordinates, as in Fig. 14.20. A *horizontal pattern* may also be plotted for more complicated antenna systems and shows the variation of field intensity with ϕ. The horizontal pattern of the current element is a circle centered at the origin since the field is not a function of the azimuth angle.

In order to obtain a quantitative expression for the power radiated, we need to apply the Poynting vector $\mathcal{P} = \mathbf{E} \times \mathbf{H}$ developed in Sec. 11.3. The instantaneous expressions for the radiation components of the electric and magnetic field intensities are

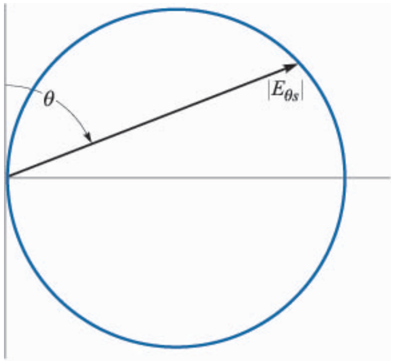


FIGURE 14.20
The polar plot of the vertical pattern of a vertical current element. The crest amplitude of $E_{\theta s}$ is plotted as a function of the polar angle θ at a constant distance r. The locus is a circle.

$$E_\theta = \eta H_\phi$$

$$H_\phi = -\frac{I_0 d}{2\lambda r}\sin\theta \sin\left(\omega t - \frac{2\pi r}{\lambda}\right)$$

and thus

$$\mathcal{P}_r = E_\theta H_\phi = \left(\frac{I_0 d}{2\lambda r}\right)^2 \eta \sin^2\theta \, \sin^2\left(\omega t - \frac{2\pi r}{\lambda}\right)$$

The total (in space) instantaneous (in time) power crossing the surface of a sphere of radius r_0 is then

$$P = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} \mathcal{P}_r r_0^2 \sin\theta d\theta d\phi$$
$$= \left(\frac{I_0 d}{\lambda}\right)^2 \eta \frac{2\pi}{3} \sin^2\left(\omega t - \frac{2\pi r_0}{\lambda}\right)$$

and the time-average power is given by one-half the maximum amplitude,

$$P_{av} = \left(\frac{I_0 d}{\lambda}\right)^2 \eta \frac{\pi}{3} = 40\pi^2 \left(\frac{I_0 d}{\lambda}\right)^2$$

where $\eta = 120\pi\,\Omega$ in free space.

This is the same power as that which would be dissipated in a resistance R_{rad} by the current I_0 in the absence of any radiation, where

$$\boxed{P_{av} = \frac{1}{2} I_0^2 R_{rad}}$$

$$R_{rad} = \frac{2P_{av}}{I_0^2} = 80\pi^2\left(\frac{d}{\lambda}\right)^2 \tag{86}$$

If we assume the differential length is 0.01λ, then R_{rad} is about 0.08 Ω. This small resistance is probably comparable to the *ohmic* resistance of a practical antenna, and thus the efficiency of the antenna might be unsatisfactorily low. Effective matching to the source also becomes very difficult to achieve, for the input reactance of a short antenna is much greater in magnitude than the input resistance R_{rad}. This is the basis for the statement that an effective antenna should be an appreciable fraction of a wavelength long.

The actual current distribution on a thin linear antenna is very nearly sinusoidal, even for antennas that may be several wavelengths long. Note that if the conductors of an open-circuited two-wire transmission line are folded back 90°, the standing-wave distribution on the line is sinusoidal. The current is zero at each end and maximizes one-quarter wavelength from each end, and the current continues to vary in this manner toward the center. The current at the center, therefore, will be very small for an antenna whose length is an integral number of wavelengths, but it will be equal to the maximum found at any point on the antenna if the antenna length is $\lambda/2$, $3\lambda/2$, $5\lambda/2$, and so forth.

On a short antenna, then, we see only the first portion of the sine wave; the amplitude of the current is zero at each end and increases approximately in a linear manner to a maximum value of I_0 at the center. This is suggested by the sketch of Fig. 14.21. Note that this antenna has identical currents in the two halves and it may be fed conveniently by a two-wire line, where the currents in the two conductors are equal in amplitude but opposite in direction. The gap at

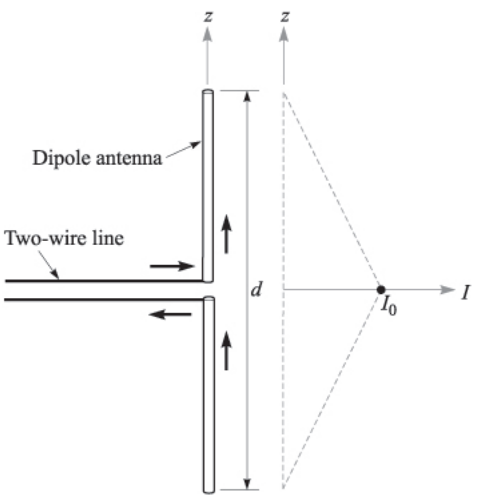


FIGURE 14.21
A short antenna ($d < \lambda/4$) has a linear current distribution and may be driven by a two-wire line.

the feed point is small and has negligible effects. A symmetrical antenna of this type is called a *dipole*. The linear current variation with distance is a reasonable assumption for antennas having an overall length less than about one-quarter wavelength.

It is possible to extend the analysis of the differential current element to the short dipole if we assume that the length is short enough that retardation effects may be neglected. That is, we consider that signals arriving at any field point P from the two ends of the antenna are in phase. The average current along the antenna is $I_0/2$, where I_0 is the input current at the center terminals. Thus, the electric and magnetic field intensities will be one-half the values given in (84) and (85), and there are no changes in the vertical and horizontal patterns. The power will be one-quarter of its previous value, and thus the radiation resistance will also be one-quarter of the value given by (86).

If we try to improve our results by assuming a sinusoidal variation of current amplitude with distance along the antenna, and if the effects of retardation are included in the analysis, then the integrations required to find **A** and the power radiated become enormously more difficult. Since we want to hurry along toward the last page, let us merely note that for the world's most popular antenna, the half-wave dipole ($d = \lambda/2$), the following results are eventually obtained:

$$E_{\theta s} = \frac{I_0 \eta}{2\pi r} \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \tag{87}$$

$$H_{\phi s} = \frac{E_{\theta s}}{\eta} \tag{88}$$

$$R_{rad} = 30\left[\frac{(2\pi)^2}{2 \cdot 2!} - \frac{(2\pi)^4}{4 \cdot 4!} + \frac{(2\pi)^6}{6 \cdot 6!} - \frac{(2\pi)^8}{8 \cdot 8!} + \ldots\right] = 73.1\ \Omega \tag{89}$$

Let us compare this accurate value with results obtained by more approximate means. Suppose we first attempt to find the radiation resistance by assuming a uniform current distribution and neglecting the effects of retardation. The result is obtained from (86) with $d/\lambda = 1/2$; $R_{rad} = 20\pi^2 = 197.4\,\Omega$. This is much greater than $73.1\,\Omega$, but we have also assumed a much greater current on the antenna than is actually present.

The result may be improved by considering a linear current distribution while still ignoring retardation. The average current is half the maximum value, the power is one-quarter, and the radiation resistance drops to $5\pi^2$ or $49.3\,\Omega$. Now the result is too small, primarily because the average value of a triangular wave is less than the average value of a sine wave.

Finally, if we assume a sinusoidal current distribution, we have an average value of $2/\pi$ times the maximum, and the radiation resistance comes to $(2/\pi)^2(20\pi^2)$, or $80\,\Omega$. This is reasonably close to the true value, and the discrepancy lies in neglecting retardation. In a linear antenna, the effect of retardation

is always one of cancellation, and therefore its consideration must always lead to smaller values of radiation resistance. This decrease is of relatively small magnitude here (from 80 to 73.1 Ω) because the current elements tending to cancel each other are those at the ends of the dipole, and these are of small amplitude; moreover, the cancellation is greatest in a direction along the antenna axis where all radiation fields are zero for a linear antenna.

Familiar antennas that fall into the dipole classification are the elements used in the common TV and FM receiving antennas.

As a final example of a practical antenna, let us collect a few facts about the *monopole* antenna. This is one-half a dipole plus a perfectly conducting plane, as shown in Fig. 14.22*a*. The image principle discussed in Sec. 5.5 provides the image shown in Fig. 14.22*b* and assures us that the fields above the plane are the same for the monopole and the dipole. Hence, the expressions of (84) and (85) are equally valid for the monopole. The Poynting vector is therefore also the same above the plane, but the integration to find the total power radiated is extended through but one-half the volume. Thus the power radiated and the radiation resistance for the monopole are half the corresponding values for the dipole. As an example, a monopole with an assumed uniform current distribution has $R_{rad} = 40\pi^2(d/\lambda)^2$; a triangular current leads to $R_{rad} = 10\pi^2(d/\lambda)^2$; and the sinusoidal current distribution of a $\lambda/4$ monopole leads to $R_{rad} = 36.5\,\Omega$.

Monopole antennas may be driven by a coaxial cable below the plane, having its center conductor connected to the antenna through a small hole, and having its outer conductor connected to the plane. If the region below the plane is inaccessible or inconvenient, the coax may be laid on top of the plane and its outer conductor connected to it. Examples of this type of antenna include AM broadcasting towers and CB antennas.

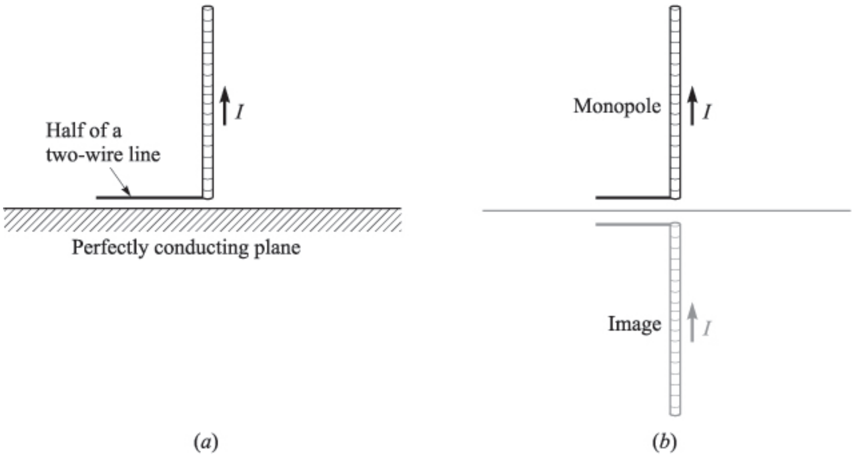


FIGURE 14.22
(*a*) An ideal monopole is always associated with a perfectly conducting plane. (*b*) The monopole plus its image form a dipole.

✔ **D14.8.** Calculate values for the curve shown in Fig. 14.19*a* at $r = 1$, 1.2, 1.4, 1.6, 1.8, and 2.

Ans. 1.00; 5.19; 2.23; −2.62; −3.22; 0.25.

✔ **D14.9.** A short antenna with a uniform current distribution in air has $I_0 d = 3 \times 10^{-4}$ A · m and $\lambda = 10$ cm. Find $|E_{\theta s}|$ at $\theta = 90°$, $\phi = 0°$, and $r =$: (a) 2 cm; (b) 20 cm; (c) 200 cm.

Ans. 24.8 V/m; 2.82 V/m; 0.283 V/m.

✔ **D14.10.** The monopole antenna of Fig. 14.22*a* has a length $d/2 = 0.080$ m and may be assumed to carry a triangular current distribution for which the feed current I_0 is 16.0 A at a frequency of 375 MHz in free space. At point P ($r = 400$ m, $\theta = 60°$, $\phi = 45°$) find: (a) $H_{\phi s}$; (b) $E_{\theta s}$; (c) the amplitude of $\mathcal{P}_r$.

Ans. j1.73 mA/m; j0.653 V/m; 1.13 mW/m^2

SUGGESTED REFERENCES

1. Ramo, S., J. R. Whinnery, and T. Van Duzer: "Fields and Waves in Communication Electronics," 3d ed., John Wiley and Sons, New York, 1990. In-depth treatment of parallel-plate and rectangular waveguides is presented in Chapter 8.
2. Marcuse, D.: "Theory of Dielectric Optical Waveguides," 2d ed., Academic Press, New York, 1990. This book provides a very general and complete discussion of dielectric slab waveguides, plus other types.
3. Buck, J. A.: "Fundamentals of Optical Fibers," Wiley-Interscience, New York, 1995. Symmetric slab dielectric guides are emphasized in this book by one of the authors.
4. Weeks, W. L.: "Antenna Engineering," McGraw-Hill Book Company, New York, 1968. This excellent text probably contains more information about antennas than you want to know.
5. Jordan, E. C. and K. G. Balmain: "Electromagnetic Waves and Radiating Systems," 2d ed., Prentice-Hall, Inc., Englewood Cliffs, N.J., 1968. This classic text provides an excellent treatment of waveguides and antennas.

PROBLEMS

14.1 A parallel-plate waveguide is known to have a cutoff wavelength for the $m = 1$ TE and TM modes of $\lambda_{c1} = 0.4$ cm. The guide is operated at wavelength $\lambda = 1$ mm. How many modes propagate?

14.2 A parallel-plate guide is to be constructed for operation in the TEM mode only over the frequency range $0 < f < 3$ GHz. The dielectric between plates is to be teflon ($\epsilon'_R = 2.1$). Determine the maximum allowable plate separation, d.

14.3 A lossless parallel-plate waveguide is known to propagate the $m = 2$ TE and TM modes at frequencies as low as 10 GHz. If the plate separation is 1 cm, determine the dielectric constant of the medium between plates.

14.4 A $d = 1$ cm parallel-plate guide is made with glass ($n = 1.45$) between plates. If the operating frequency is 32 GHz, which modes will propagate?

14.5 For the guide of Problem 14.4, and at the 32 GHz frequency, determine the difference between the group delays of the highest-order mode (TE or TM) and the TEM mode. Assume a propagation distance of 10 cm.

14.6 The cutoff frequency of the $m = 1$ TE and TM modes in a parallel-plate guide is known to be $f_{c1} = 7.5$ GHz. The guide is used at wavelength $\lambda = 1.5$ cm. Find the group velocity of the $m = 2$ TE and TM modes.

14.7 A parallel-plate guide is partially filled with two lossless dielectrics (Fig. 14.23) where $\epsilon'_{R1} = 4.0$, $\epsilon'_{R2} = 2.1$, and $d = 1$ cm. At a certain frequency, it is found that the TM_1 mode propagates through the guide without suffering any reflective loss at the dielectric interface. (a) Find this frequency. (b) Is the guide operating at a single TM mode at the frequency found in part (a)? Hint: Remember Brewster's angle?

14.8 In the guide of Problem 14.7, it is found that $m = 1$ modes propagating from left to right totally reflect at the interface, so that no power is transmitted into the region of dielectric constant ϵ'_{R2}. (a) Determine the range of frequencies over which this will occur. (b) Does your part (a) answer in any way relate to the cutoff frequency for $m = 1$ modes in either region? Hint: Remember the critical angle?

14.9 A rectangular waveguide has dimensions $a = 6$ cm and $b = 4$ cm. (a) Over what range of frequencies will the guide operate single mode? (b) Over what frequency range will the guide support *both* TE_{10} and TE_{01} modes and no others?

14.10 Two rectangular waveguides are joined end-to-end. The guides have identical dimensions, where $a = 2b$. One guide is air-filled; the other is filled with a lossless dielectric characterized by ϵ'_R. (a) Determine the maximum allowable value of ϵ'_R such that single-mode operation can be simultaneously ensured in *both* guides at some frequency. (b) Write an expression for the frequency range over which single-mode operation will occur in both guides; your answer should be in terms of ϵ'_R, guide dimensions as needed, and other known constants.

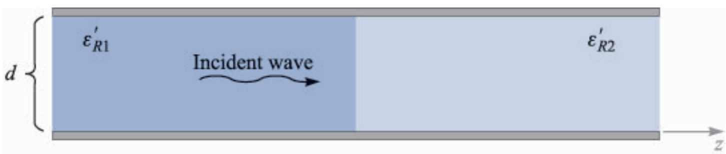


FIGURE 14.23
See Problem 14.7.

14.11 An air-filled rectangular waveguide is to be constructed for single-mode operation at 15 GHz. Specify the guide dimensions, a and b, such that the design frequency is 10% higher than the cutoff frequency for the TE_{10} mode, while being 10% lower than the cutoff frequency for the next higher-order mode.

14.12 Using the relation $\mathcal{P}_{av} = \frac{1}{2}\text{Re}\{\mathbf{E}_s \times \mathbf{H}_s^*\}$, and Eqs. (44) through (46), show that the average power density in the TE_{10} mode in a rectangular waveguide is given by

$$\mathcal{P}_{av} = \frac{\beta_{10}}{2\omega\mu} E_0^2 \sin(\kappa_{10} x)\mathbf{a}_z \qquad \text{W/m}^2$$

14.13 Integrate the result of Problem 14.12 over the guide cross section, $0 < x < a$, $0 < y < b$, to show that the power in Watts transmitted down the guide is given as

$$P = \frac{\beta_{10} ab}{4\omega\mu} E_0^2 = \frac{ab}{4\eta} E_0^2 \sin\theta_{10} \qquad \text{W}$$

where $\eta = \sqrt{\mu\epsilon}$ and θ_{10} is the wave angle associated with the TE_{10} mode. Interpret.

14.14 Show that the group dispersion parameter, $d^2\beta/d\omega^2$, for a given mode in a parallel-plate or rectangular waveguide is given by

$$\frac{d^2\beta}{d\omega^2} = -\frac{n}{\omega c}\left(\frac{\omega_c}{\omega}\right)^2 \left[1 - \left(\frac{\omega_c}{\omega}\right)^2\right]^{-3/2}$$

where ω_c is the radian cutoff frequency for the mode in question (note that the first derivative form was already found, resulting in Eq. (23)).

14.15 Consider a transform-limited pulse of center frequency $f = 10$ GHz, and of full-width $2T = 1.0$ ns. The pulse propagates in a lossless single mode rectangular guide which is air-filled and in which the 10 GHz operating frequency is 1.1 times the cutoff frequency of the TE_{10} mode. Using the result of Problem 14.14, determine the length of guide over which the pulse broadens to twice its initial width. What simple step can be taken to reduce the amount of pulse broadening in this guide, while maintaining the same initial pulse width? Additional background for this problem is found in Sec. 12.6.

14.16 A symmetric dielectric slab waveguide has a slab thickness $d = 10\,\mu$m, with $n_1 = 1.48$ and $n_2 = 1.45$. If the operating wavelength is $\lambda = 1.3\,\mu$m, what modes will propagate?

14.17 A symmetric slab waveguide is known to support only a single pair of TE and TM modes at wavelength $\lambda = 1.55\,\mu$m. If the slab thickness is 5 μm, what is the maximum value of n_1 if $n_2 = 3.3$?

14.18 $n_1 = 1.45$, $n_2 = 1.50$, and $d = 10\,\mu$m in a symmetric slab waveguide. (a) What is the phase velocity of the $m = 1$ TE or TM mode at cutoff? (b) What is the phase velocity of the $m = 2$ TE or TM modes at cutoff?

14.19 An *asymmetric* slab waveguide is shown in Fig. 14.24. In this case, the regions above and below the slab have unequal refractive indices, where $n_1 < n_3 < n_2$. (a) Write, in terms of the appropriate indices, an expression for the minimum possible wave angle, θ_1, that a guided mode may have. (b) Write an expression for the maximum phase velocity a guided mode may have in this structure, using given or known parameters.

14.20 A step index optical fiber is known to be single mode at wavelengths $\lambda > 1.2\,\mu$m. Another fiber is to be fabricated from the same materials, but it is to be single mode at wavelengths $\lambda > 0.63\,\mu$m. By what percentage must the core radius of the new fiber differ from the old one, and should it be larger or smaller?

14.21 A short dipole carrying current $I_0 \cos\omega t$ in the $\mathbf{a}_z$ direction is located at the origin in free space. (a) If $\beta = 1$ rad/m, $r = 2$ m, $\theta = 45^\circ$, $\phi = 0$, and $t = 0$, give a unit vector in rectangular components that shows the instantaneous direction of $\mathbf{E}$. (b) What fraction of the total average power is radiated in the belt, $80^\circ < \theta < 100^\circ$?

14.22 Prepare a curve, r vs. θ in polar coordinates, showing the locus in the $\phi = 0$ plane where: (a) the radiation field $|E_{\theta s}|$ is one-half of its value at $r = 10^4$ m, $\theta = \pi/2$; (b) the average radiated power density $\mathcal{P}_{r,av}$ is one-half of its value at $r = 10^4$ m, $\theta = \pi/2$.

14.23 Two short antennas at the origin in free space carry identical currents of $5\cos\omega t$ A, one in the $\mathbf{a}_z$ direction, one in the $\mathbf{a}_y$ direction. Let $\lambda = 2\pi$ m and $d = 0.1$ m. Find $\mathbf{E}_s$ at the distant point: (a) $(x = 0, y = 1000, z = 0)$; (b) (0,0,1000); (c) (1000,0,0); (d) Find $\mathbf{E}$ at (1000,0,0) at $t = 0$. (e) Find $|\mathbf{E}|$ at (1000,0,0) at $t = 0$.

14.24 A short current element has $d = 0.03\lambda$. Calculate the radiation resistance for each of the following current distributions: (a) uniform, I_0; (b) linear, $I(z) = I_0(0.5d - |z|)/0.5d$; (c) step, I_0 for $0 < |z| < 0.25d$ and $0.5I_0$ for $0.25d < |z| < 0.5d$.

14.25 A dipole antenna in free space has a linear current distribution. If the length d is 0.02λ, what value of I_0 is required to: (a) provide a radiation-field amplitude of 100 mV/m at a distance of 1 mi, at $\theta = 90^\circ$; (b) radiate a total power of 1 W?

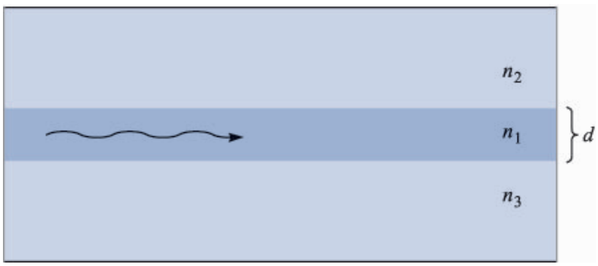


FIGURE 14.24
See Problem 14.19.

14.26 A monopole antenna in free space, extending vertically over a perfectly conducting plane, has a linear current distribution. If the length of the antenna is 0.01λ, what value of I_0 is required to: (a) provide a radiation-field amplitude of 100 mV/m at a distance of 1 mi, at $\theta = 90^\circ$; (b) radiate a total power of 1 W?

14.27 The radiation field of a certain short vertical current element is $E_{\theta s} = (20/r)\sin\theta\, e^{-j10\pi r}$ V/m if it is located at the origin in free space. (a) Find $E_{\theta s}$ at $P(r = 100,\ \theta = 90^\circ,\ \phi = 30^\circ)$. (b) Find $E_{\theta s}$ at $P(100, 90^\circ, 30^\circ)$ if the vertical element is located at $A(0.1, 90^\circ, 90^\circ)$. (c) Find $E_{\theta s}$ at $P(100, 90^\circ, 30^\circ)$ if identical vertical elements are located at $A(0.1, 90^\circ, 90^\circ)$ and $B(0.1, 90^\circ, 270^\circ)$.

Problem Solutions for E-text

1.3. The vector from the origin to the point A is given as $(6, -2, -4)$, and the unit vector directed from the origin toward point B is $(2, -2, 1)/3$. If points A and B are ten units apart, find the coordinates of point B.

With $\mathbf{A} = (6, -2, -4)$ and $\mathbf{B} = \frac{1}{3}B(2, -2, 1)$, we use the fact that $|\mathbf{B} - \mathbf{A}| = 10$, or
$|(6 - \frac{2}{3}B)\mathbf{a}_x - (2 - \frac{2}{3}B)\mathbf{a}_y - (4 + \frac{1}{3}B)\mathbf{a}_z| = 10$
Expanding, obtain
$36 - 8B + \frac{4}{9}B^2 + 4 - \frac{8}{3}B + \frac{4}{9}B^2 + 16 + \frac{8}{3}B + \frac{1}{9}B^2 = 100$
or $B^2 - 8B - 44 = 0$. Thus $B = \frac{8 \pm \sqrt{64-176}}{2} = 11.75$ (taking positive option) and so

$$\mathbf{B} = \frac{2}{3}(11.75)\mathbf{a}_x - \frac{2}{3}(11.75)\mathbf{a}_y + \frac{1}{3}(11.75)\mathbf{a}_z = \underline{7.83\mathbf{a}_x - 7.83\mathbf{a}_y + 3.92\mathbf{a}_z}$$

1.17. Point $A(-4, 2, 5)$ and the two vectors, $\mathbf{R}_{AM} = (20, 18, -10)$ and $\mathbf{R}_{AN} = (-10, 8, 15)$, define a triangle.

a) Find a unit vector perpendicular to the triangle: Use

$$\mathbf{a}_p = \frac{\mathbf{R}_{AM} \times \mathbf{R}_{AN}}{|\mathbf{R}_{AM} \times \mathbf{R}_{AN}|} = \frac{(350, -200, 340)}{527.35} = \underline{(0.664, -0.379, 0.645)}$$

The vector in the opposite direction to this one is also a valid answer.

b) Find a unit vector in the plane of the triangle and perpendicular to $\mathbf{R}_{AN}$:

$$\mathbf{a}_{AN} = \frac{(-10, 8, 15)}{\sqrt{389}} = (-0.507, 0.406, 0.761)$$

Then

$$\mathbf{a}_{pAN} = \mathbf{a}_p \times \mathbf{a}_{AN} = (0.664, -0.379, 0.645) \times (-0.507, 0.406, 0.761) = \underline{(-0.550, -0.832, 0.077)}$$

The vector in the opposite direction to this one is also a valid answer.

c) Find a unit vector in the plane of the triangle that bisects the interior angle at A: A non-unit vector in the required direction is $(1/2)(\mathbf{a}_{AM} + \mathbf{a}_{AN})$, where

$$\mathbf{a}_{AM} = \frac{(20, 18, -10)}{|(20, 18, -10)|} = (0.697, 0.627, -0.348)$$

Now

$$\frac{1}{2}(\mathbf{a}_{AM} + \mathbf{a}_{AN}) = \frac{1}{2}[(0.697, 0.627, -0.348) + (-0.507, 0.406, 0.761)] = (0.095, 0.516, 0.207)$$

Finally,

$$\mathbf{a}_{bis} = \frac{(0.095, 0.516, 0.207)}{|(0.095, 0.516, 0.207)|} = \underline{(0.168, 0.915, 0.367)}$$

1.27. The surfaces $r = 2$ and 4, $\theta = 30°$ and $50°$, and $\phi = 20°$ and $60°$ identify a closed surface.

a) Find the enclosed volume: This will be

$$\text{Vol} = \int_{20°}^{60°} \int_{30°}^{50°} \int_{2}^{4} r^2 \sin\theta dr d\theta d\phi = \underline{2.91}$$

where degrees have been converted to radians.

b) Find the total area of the enclosing surface:

$$\text{Area} = \int_{20°}^{60°} \int_{30°}^{50°} (4^2 + 2^2) \sin\theta d\theta d\phi + \int_{2}^{4} \int_{20°}^{60°} r(\sin 30° + \sin 50°) dr d\phi + 2 \int_{30°}^{50°} \int_{2}^{4} r dr d\theta = \underline{12.61}$$

c) Find the total length of the twelve edges of the surface:

$$\text{Length} = 4 \int_{2}^{4} dr \ + \ 2 \int_{30°}^{50°} (4+2) d\theta + \int_{20°}^{60°} (4\sin 50° + 4\sin 30° + 2\sin 50° + 2\sin 30°) d\phi = \underline{17.49}$$

d) Find the length of the longest straight line that lies entirely within the surface: This will be from $A(r = 2, \theta = 50°, \phi = 20°)$ to $B(r = 4, \theta = 30°, \phi = 60°)$ or

$$A(x = 2\sin 50° \cos 20°, y = 2\sin 50° \sin 20°, z = 2\cos 50°)$$

to

$$B(x = 4\sin 30° \cos 60°, y = 4\sin 30° \sin 60°, z = 4\cos 30°)$$

or finally $A(1.44, 0.52, 1.29)$ to $B(1.00, 1.73, 3.46)$. Thus $\mathbf{B} - \mathbf{A} = (-0.44, 1.21, 2.18)$ and

$$\text{Length} = |\mathbf{B} - \mathbf{A}| = \underline{2.53}$$

2.5. Let a point charge $Q_1 25$ nC be located at $P_1(4, -2, 7)$ and a charge $Q_2 = 60$ nC be at $P_2(-3, 4, -2)$.

a) If $\epsilon = \epsilon_0$, find $\mathbf{E}$ at $P_3(1, 2, 3)$: This field will be

$$\mathbf{E} = \frac{10^{-9}}{4\pi\epsilon_0} \left[\frac{25\mathbf{R}_{13}}{|\mathbf{R}_{13}|^3} + \frac{60\mathbf{R}_{23}}{|\mathbf{R}_{23}|^3} \right]$$

where $\mathbf{R}_{13} = -3\mathbf{a}_x + 4\mathbf{a}_y - 4\mathbf{a}_z$ and $\mathbf{R}_{23} = 4\mathbf{a}_x - 2\mathbf{a}_y + 5\mathbf{a}_z$. Also, $|\mathbf{R}_{13}| = \sqrt{41}$ and $|\mathbf{R}_{23}| = \sqrt{45}$. So

$$\mathbf{E} = \frac{10^{-9}}{4\pi\epsilon_0} \left[\frac{25 \times (-3\mathbf{a}_x + 4\mathbf{a}_y - 4\mathbf{a}_z)}{(41)^{1.5}} + \frac{60 \times (4\mathbf{a}_x - 2\mathbf{a}_y + 5\mathbf{a}_z)}{(45)^{1.5}} \right] = \underline{4.58\mathbf{a}_x - 0.15\mathbf{a}_y + 5.51\mathbf{a}_z}$$

b) At what point on the y axis is $E_x = 0$? P_3 is now at $(0, y, 0)$, so $\mathbf{R}_{13} = -4\mathbf{a}_x + (y+2)\mathbf{a}_y - 7\mathbf{a}_z$ and $\mathbf{R}_{23} = 3\mathbf{a}_x + (y-4)\mathbf{a}_y + 2\mathbf{a}_z$. Also, $|\mathbf{R}_{13}| = \sqrt{65 + (y+2)^2}$ and $|\mathbf{R}_{23}| = \sqrt{13 + (y-4)^2}$. Now the x component of $\mathbf{E}$ at the new P_3 will be:

$$E_x = \frac{10^{-9}}{4\pi\epsilon_0} \left[\frac{25 \times (-4)}{[65 + (y+2)^2]^{1.5}} + \frac{60 \times 3}{[13 + (y-4)^2]^{1.5}} \right]$$

To obtain $E_x = 0$, we require the expression in the large brackets to be zero. This expression simplifies to the following quadratic:

$$0.48y^2 + 13.92y + 73.10 = 0$$

which yields the two values: $y = \underline{-6.89, -22.11}$

2.19. A uniform line charge of 2 μC/m is located on the z axis. Find $\mathbf{E}$ in cartesian coordinates at $P(1, 2, 3)$ if the charge extends from

a) $-\infty < z < \infty$: With the infinite line, we know that the field will have only a radial component in cylindrical coordinates (or x and y components in cartesian). The field from an infinite line on the z axis is generally $\mathbf{E} = [\rho_l/(2\pi\epsilon_0\rho)]\mathbf{a}_\rho$. Therefore, at point P:

$$\mathbf{E}_P = \frac{\rho_l}{2\pi\epsilon_0}\frac{\mathbf{R}_{zP}}{|\mathbf{R}_{zP}|^2} = \frac{(2\times 10^{-6})}{2\pi\epsilon_0}\frac{\mathbf{a}_x + 2\mathbf{a}_y}{5} = \underline{7.2\mathbf{a}_x + 14.4\mathbf{a}_y \ \text{kV/m}}$$

where $\mathbf{R}_{zP}$ is the vector that extends from the line charge to point P, and is perpendicular to the z axis; i.e., $\mathbf{R}_{zP} = (1, 2, 3) - (0, 0, 3) = (1, 2, 0)$.

b) $-4 \le z \le 4$: Here we use the general relation

$$\mathbf{E}_P = \int \frac{\rho_l dz}{4\pi\epsilon_0}\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|^3}$$

where $\mathbf{r} = \mathbf{a}_x + 2\mathbf{a}_y + 3\mathbf{a}_z$ and $\mathbf{r}' = z\mathbf{a}_z$. So the integral becomes

$$\mathbf{E}_P = \frac{(2\times 10^{-6})}{4\pi\epsilon_0}\int_{-4}^{4}\frac{\mathbf{a}_x + 2\mathbf{a}_y + (3-z)\mathbf{a}_z}{[5+(3-z)^2]^{1.5}}\,dz$$

Using integral tables, we obtain:

$$\mathbf{E}_P = 3597\left[\frac{(\mathbf{a}_x + 2\mathbf{a}_y)(z-3) + 5\mathbf{a}_z}{(z^2 - 6z + 14)}\right]_{-4}^{4} \ \text{V/m} = \underline{4.9\mathbf{a}_x + 9.8\mathbf{a}_y + 4.9\mathbf{a}_z \ \text{kV/m}}$$

The student is invited to verify that when evaluating the above expression over the limits $-\infty < z < \infty$, the z component vanishes and the x and y components become those found in part a.

2.27. Given the electric field $\mathbf{E} = (4x - 2y)\mathbf{a}_x - (2x + 4y)\mathbf{a}_y$, find:

a) the equation of the streamline that passes through the point $P(2, 3, -4)$: We write

$$\frac{dy}{dx} = \frac{E_y}{E_x} = \frac{-(2x+4y)}{(4x-2y)}$$

Thus

$$2(x\,dy + y\,dx) = y\,dy - x\,dx$$

or

$$2\,d(xy) = \frac{1}{2}\,d(y^2) - \frac{1}{2}\,d(x^2)$$

So

$$C_1 + 2xy = \frac{1}{2}y^2 - \frac{1}{2}x^2$$

or

$$y^2 - x^2 = 4xy + C_2$$

Evaluating at $P(2, 3, -4)$, obtain:

$$9 - 4 = 24 + C_2, \text{ or } C_2 = -19$$

Finally, at P, the requested equation is

$$\underline{y^2 - x^2 = 4xy - 19}$$

b) a unit vector specifying the direction of $\mathbf{E}$ at $Q(3, -2, 5)$: Have $\mathbf{E}_Q = [4(3) + 2(2)]\mathbf{a}_x - [2(3) - 4(2)]\mathbf{a}_y = 16\mathbf{a}_x + 2\mathbf{a}_y$. Then $|\mathbf{E}| = \sqrt{16^2 + 4} = 16.12$ So

$$\mathbf{a}_Q = \frac{16\mathbf{a}_x + 2\mathbf{a}_y}{16.12} = \underline{0.99\mathbf{a}_x + 0.12\mathbf{a}_y}$$

3.5. Let $\mathbf{D} = 4xy\mathbf{a}_x + 2(x^2 + z^2)\mathbf{a}_y + 4yz\mathbf{a}_z$ C/m^2 and evaluate surface integrals to find the total charge enclosed in the rectangular parallelepiped $0 < x < 2$, $0 < y < 3$, $0 < z < 5$ m: Of the 6 surfaces to consider, only 2 will contribute to the net outward flux. Why? First consider the planes at $y = 0$ and 3. The y component of $\mathbf{D}$ will penetrate those surfaces, but will be inward at $y = 0$ and outward at $y = 3$, while having the same magnitude in both cases. These fluxes will thus cancel. At the $x = 0$ plane, $D_x = 0$ and at the $z = 0$ plane, $D_z = 0$, so there will be no flux contributions from these surfaces. This leaves the 2 remaining surfaces at $x = 2$ and $z = 5$. The net outward flux becomes:

$$\begin{aligned}\Phi &= \int_0^5 \int_0^3 \mathbf{D}\big|_{x=2} \cdot \mathbf{a}_x \, dy\, dz + \int_0^3 \int_0^2 \mathbf{D}\big|_{z=5} \cdot \mathbf{a}_z \, dx\, dy \\ &= 5\int_0^3 4(2)y\, dy \; + \; 2\int_0^3 4(5)y\, dy = \underline{360 \text{ C}}\end{aligned}$$

3.21. Calculate the divergence of $\mathbf{D}$ at the point specified if

a) $\mathbf{D} = (1/z^2)\left[10xyz\,\mathbf{a}_x + 5x^2z\,\mathbf{a}_y + (2z^3 - 5x^2y)\,\mathbf{a}_z\right]$ at $P(-2, 3, 5)$: We find

$$\nabla \cdot \mathbf{D} = \left[\frac{10y}{z} + 0 + 2 + \frac{10x^2y}{z^3}\right]_{(-2,3,5)} = \underline{8.96}$$

b) $\mathbf{D} = 5z^2\mathbf{a}_\rho + 10\rho z\,\mathbf{a}_z$ at $P(3, -45°, 5)$: In cylindrical coordinates, we have

$$\nabla \cdot \mathbf{D} = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho D_\rho) + \frac{1}{\rho}\frac{\partial D_\phi}{\partial \phi} + \frac{\partial D_z}{\partial z} = \left[\frac{5z^2}{\rho} + 10\rho\right]_{(3,-45°,5)} = \underline{71.67}$$

c) $\mathbf{D} = 2r\sin\theta\sin\phi\,\mathbf{a}_r + r\cos\theta\sin\phi\,\mathbf{a}_\theta + r\cos\phi\,\mathbf{a}_\phi$ at $P(3, 45°, -45°)$: In spherical coordinates, we have

$$\begin{aligned}\nabla \cdot \mathbf{D} &= \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 D_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta D_\theta) + \frac{1}{r\sin\theta}\frac{\partial D_\phi}{\partial\phi} \\ &= \left[6\sin\theta\sin\phi + \frac{\cos 2\theta\sin\phi}{\sin\theta} - \frac{\sin\phi}{\sin\theta}\right]_{(3,45°,-45°)} = \underline{-2}\end{aligned}$$

3.27. Let $\mathbf{D} = 5r^2\mathbf{a}_r$ mC/m^2 for $r < 0.08$ m and $\mathbf{D} = 0.1\,\mathbf{a}_r/r^2$ mC/m^2 for $r > 0.08$ m.

a) Find ρ_v for $r = 0.06$ m: This radius lies within the first region, and so

$$\rho_v = \nabla \cdot \mathbf{D} = \frac{1}{r^2}\frac{d}{dr}(r^2 D_r) = \frac{1}{r^2}\frac{d}{dr}(5r^4) = 20r \text{ C/m}^3$$

which when evaluated at $r = 0.06$ yields $\rho_v(r = .06) = \underline{1.20 \text{ mC/m}^3}$.

b) Find ρ_v for $r = 0.1$ m: This is in the region where the second field expression is valid. The $1/r^2$ dependence of this field yields a zero divergence (shown in Problem 3.23), and so the volume charge density is zero at 0.1 m.

c) What surface charge density could be located at $r = 0.08$ m to cause $\mathbf{D} = 0$ for $r > 0.08$ m? The total surface charge should be equal and opposite to the total volume charge. The latter is

$$Q = \int_0^{2\pi}\int_0^{\pi}\int_0^{.08} 20r(\text{mC/m}^3)\, r^2 \sin\theta\, dr\, d\theta\, d\phi = 2.57 \times 10^{-3} \text{ mC} = 2.57\,\mu\text{C}$$

So now

$$\rho_s = -\left[\frac{2.57}{4\pi(.08)^2}\right] = \underline{-32\,\mu\text{C/m}^2}$$

4.5. Compute the value of $\int_A^P \mathbf{G} \cdot d\mathbf{L}$ for $\mathbf{G} = 2y\mathbf{a}_x$ with $A(1, -1, 2)$ and $P(2, 1, 2)$ using the path:

a) straight-line segments $A(1, -1, 2)$ to $B(1, 1, 2)$ to $P(2, 1, 2)$: In general we would have

$$\int_A^P \mathbf{G} \cdot d\mathbf{L} = \int_A^P 2y\, dx$$

The change in x occurs when moving between B and P, during which $y = 1$. Thus

$$\int_A^P \mathbf{G} \cdot d\mathbf{L} = \int_B^P 2y\, dx = \int_1^2 2(1)dx = \underline{2}$$

b) straight-line segments $A(1, -1, 2)$ to $C(2, -1, 2)$ to $P(2, 1, 2)$: In this case the change in x occurs when moving from A to C, during which $y = -1$. Thus

$$\int_A^P \mathbf{G} \cdot d\mathbf{L} = \int_A^C 2y\, dx = \int_1^2 2(-1)dx = \underline{-2}$$

4.17. Uniform surface charge densities of 6 and 2 nC/m^2 are present at $\rho = 2$ and 6 cm respectively, in free space. Assume $V = 0$ at $\rho = 4$ cm, and calculate V at:

a) $\rho = 5$ cm: Since $V = 0$ at 4 cm, the potential at 5 cm will be the potential difference between points 5 and 4:

$$V_5 = -\int_4^5 \mathbf{E} \cdot d\mathbf{L} = -\int_4^5 \frac{a\rho_{sa}}{\epsilon_0 \rho}\, d\rho = -\frac{(.02)(6 \times 10^{-9})}{\epsilon_0} \ln\left(\frac{5}{4}\right) = \underline{-3.026\text{ V}}$$

b) $\rho = 7$ cm: Here we integrate piecewise from $\rho = 4$ to $\rho = 7$:

$$V_7 = -\int_4^6 \frac{a\rho_{sa}}{\epsilon_0 \rho} d\rho - \int_6^7 \frac{(a\rho_{sa} + b\rho_{sb})}{\epsilon_0 \rho} d\rho$$

With the given values, this becomes

$$V_7 = -\left[\frac{(.02)(6 \times 10^{-9})}{\epsilon_0}\right] \ln\left(\frac{6}{4}\right) - \left[\frac{(.02)(6 \times 10^{-9}) + (.06)(2 \times 10^{-9})}{\epsilon_0}\right] \ln\left(\frac{7}{6}\right)$$
$$= \underline{-9.678\,\text{V}}$$

4.31. (continued) The integral evaluates as follows:

$$W_E = 200\epsilon_0 \int_1^2 \int_1^2 \left[-\left(\frac{1}{3}\right)\frac{1}{x^3y^2z^2} - \frac{1}{xy^4z^2} - \frac{1}{xy^2z^4}\right]_1^2 dy\,dz$$
$$= 200\epsilon_0 \int_1^2 \int_1^2 \left[\left(\frac{7}{24}\right)\frac{1}{y^2z^2} + \left(\frac{1}{2}\right)\frac{1}{y^4z^2} + \left(\frac{1}{2}\right)\frac{1}{y^2z^4}\right] dy\,dz$$
$$= 200\epsilon_0 \int_1^2 \left[-\left(\frac{7}{24}\right)\frac{1}{yz^2} - \left(\frac{1}{6}\right)\frac{1}{y^3z^2} - \left(\frac{1}{2}\right)\frac{1}{yz^4}\right]_1^2 dz$$
$$= 200\epsilon_0 \int_1^2 \left[\left(\frac{7}{48}\right)\frac{1}{z^2} + \left(\frac{7}{48}\right)\frac{1}{z^2} + \left(\frac{1}{4}\right)\frac{1}{z^4}\right] dz$$
$$= 200\epsilon_0(3)\left[\frac{7}{96}\right] = \underline{387\,\text{pJ}}$$

b) What value would be obtained by assuming a uniform energy density equal to the value at the center of the cube? At $C(1.5, 1.5, 1.5)$ the energy density is

$$w_E = 200\epsilon_0(3)\left[\frac{1}{(1.5)^4(1.5)^2(1.5)^2}\right] = 2.07 \times 10^{-10}\ \text{J/m}^3$$

This, multiplied by a cube volume of 1, produces an energy value of $\underline{207\,\text{pJ}}$.

5.1. Given the current density $\mathbf{J} = -10^4[\sin(2x)e^{-2y}\mathbf{a}_x + \cos(2x)e^{-2y}\mathbf{a}_y]\,\text{kA/m}^2$:

a) Find the total current crossing the plane $y = 1$ in the $\mathbf{a}_y$ direction in the region $0 < x < 1$, $0 < z < 2$: This is found through

$$I = \int\int_S \mathbf{J}\cdot\mathbf{n}\Big|_S da = \int_0^2\int_0^1 \mathbf{J}\cdot\mathbf{a}_y\Big|_{y=1} dx\,dz = \int_0^2\int_0^1 -10^4\cos(2x)e^{-2}\,dx\,dz$$
$$= -10^4(2)\frac{1}{2}\sin(2x)\Big|_0^1 e^{-2} = \underline{-1.23\,\text{MA}}$$

b) Find the total current leaving the region $0 < x, x < 1$, $2 < z < 3$ by integrating $\mathbf{J}\cdot\mathbf{dS}$ over the surface of the cube: Note first that current through the top and bottom surfaces will not exist, since $\mathbf{J}$ has no z component. Also note that there will be no current through the $x = 0$ plane,

since $J_x = 0$ there. Current will pass through the three remaining surfaces, and will be found through

$$I = \int_2^3 \int_0^1 \mathbf{J} \cdot (-\mathbf{a}_y)\Big|_{y=0} dx\,dz + \int_2^3 \int_0^1 \mathbf{J} \cdot (\mathbf{a}_y)\Big|_{y=1} dx\,dz + \int_2^3 \int_0^1 \mathbf{J} \cdot (\mathbf{a}_x)\Big|_{x=1} dy\,dz$$

$$= 10^4 \int_2^3 \int_0^1 \left[\cos(2x)e^{-0} - \cos(2x)e^{-2}\right] dx\,dz - 10^4 \int_2^3 \int_0^1 \sin(2)e^{-2y}\,dy\,dz$$

$$= 10^4 \left(\frac{1}{2}\right) \sin(2x)\Big|_0^1 (3-2)\left[1 - e^{-2}\right] + 10^4 \left(\frac{1}{2}\right) \sin(2)e^{-2y}\Big|_0^1 (3-2) = \underline{0}$$

c) Repeat part *b*, but use the divergence theorem: We find the net outward current through the surface of the cube by integrating the divergence of **J** over the cube volume. We have

$$\nabla \cdot \mathbf{J} = \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} = -10^{-4}\left[2\cos(2x)e^{-2y} - 2\cos(2x)e^{-2y}\right] = \underline{0} \text{ as expected}$$

5.11. Two perfectly-conducting cylindrical surfaces are located at $\rho = 3$ and $\rho = 5$ cm. The total current passing radially outward through the medium between the cylinders is 3 A dc. *Assume the cylinders are both of length l.*

a) Find the voltage and resistance between the cylinders, and **E** in the region between the cylinders, if a conducting material having $\sigma = 0.05\,\mathrm{S/m}$ is present for $3 < \rho < 5$ cm: Given the current, and knowing that it is radially-directed, we find the current density by dividing it by the area of a cylinder of radius ρ and length l:

$$\mathbf{J} = \frac{3}{2\pi\rho l}\,\mathbf{a}_\rho \text{ A/m}^2$$

Then the electric field is found by dividing this result by σ:

$$\mathbf{E} = \frac{3}{2\pi\sigma\rho l}\,\mathbf{a}_\rho = \underline{\frac{9.55}{\rho l}\mathbf{a}_\rho \text{ V/m}}$$

The voltage between cylinders is now:

$$V = -\int_5^3 \mathbf{E} \cdot \mathbf{dL} = \int_3^5 \frac{9.55}{\rho l}\mathbf{a}_\rho \cdot \mathbf{a}_\rho d\rho = \frac{9.55}{l}\ln\left(\frac{5}{3}\right) = \underline{\frac{4.88}{l}\text{ V}}$$

Now, the resistance will be

$$R = \frac{V}{I} = \frac{4.88}{3l} = \underline{\frac{1.63}{l}\ \Omega}$$

b) Show that integrating the power dissipated per unit volume over the volume gives the total dissipated power: We calculate

$$P = \int_v \mathbf{E} \cdot \mathbf{J}\,dv = \int_0^l \int_0^{2\pi} \int_{.03}^{.05} \frac{3^2}{(2\pi)^2\rho^2(.05)l^2}\,\rho\,d\rho\,d\phi\,dz = \frac{3^2}{2\pi(.05)l}\ln\left(\frac{5}{3}\right) = \frac{14.64}{l}\text{ W}$$

We also find the power by taking the product of voltage and current:

$$P = VI = \frac{4.88}{l}(3) = \frac{14.64}{l}\text{ W}$$

which is in aggreement with the power density integration.

5.39. A parallel plate capacitor is filled with a nonuniform dielectric characterized by $\epsilon_R = 2 + 2 \times 10^6 x^2$, where x is the distance from one plate. If $S = 0.02\,\text{m}^2$, and $d = 1$ mm, find C: Start by assuming charge density ρ_s on the top plate. **D** will, as usual, be x-directed, originating at the top plate and terminating on the bottom plate. The key here is that **D** *will be constant over the distance between plates.* This can be understood by considering the x-varying dielectric as constructed of many thin layers, each having constant permittivity. The permittivity changes from layer to layer to approximate the given function of x. The approximation becomes exact as the layer thicknesses approach zero. We know that **D**, which is normal to the layers, will be continuous across each boundary, and so **D** is constant over the plate separation distance, and will be given in magnitude by ρ_s. The electric field magnitude is now

$$E = \frac{D}{\epsilon_0 \epsilon_R} = \frac{\rho_s}{\epsilon_0(2 + 2 \times 10^6 x^2)}$$

The voltage beween plates is then

$$V_0 = \int_0^{10^{-3}} \frac{\rho_s \, dx}{\epsilon_0(2 + 2 \times 10^6 x^2)} = \frac{\rho_s}{\epsilon_0} \frac{1}{\sqrt{4 \times 10^6}} \tan^{-1}\left(\frac{x\sqrt{4 \times 10^6}}{2}\right)\Big|_0^{10^{-3}} = \frac{\rho_s}{\epsilon_0} \frac{1}{2 \times 10^3} \left(\frac{\pi}{4}\right)$$

Now $Q = \rho_s(.02)$, and so

$$C = \frac{Q}{V_0} = \frac{\rho_s(.02)\epsilon_0(2 \times 10^3)(4)}{\rho_s \pi} = 4.51 \times 10^{-10}\,\text{F} = \underline{451\,\text{pF}}$$

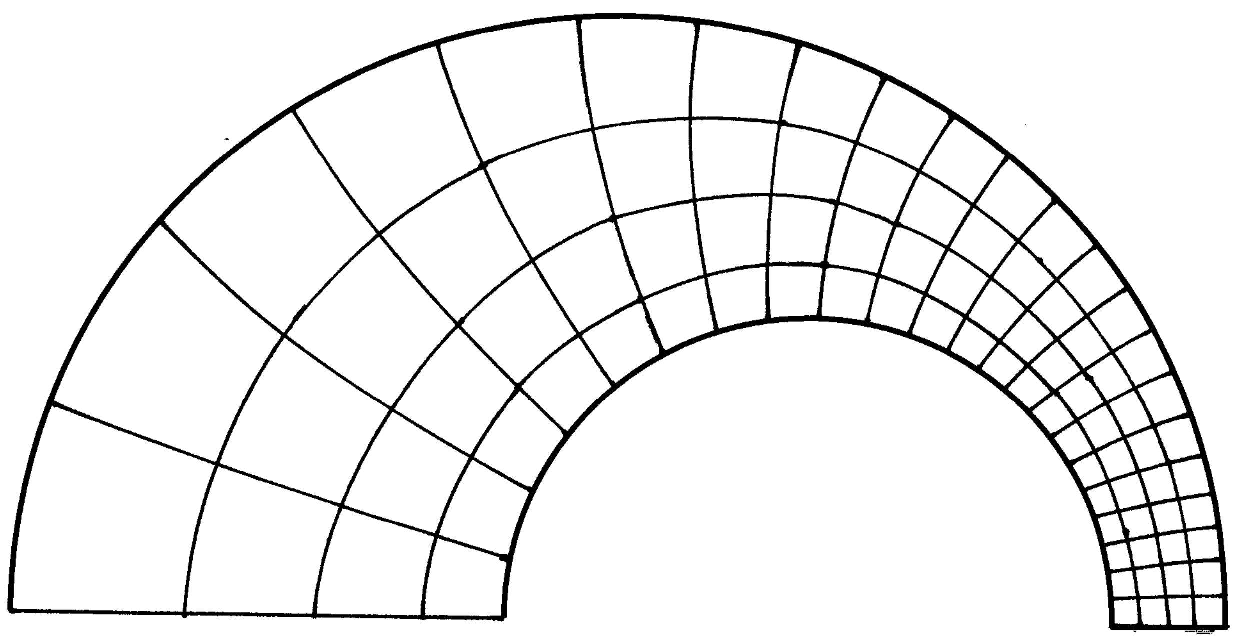
6.3. Construct a curvilinear square map of the potential field between two parallel circular cylinders, one of 4-cm radius inside one of 8-cm radius. The two axes are displaced by 2.5 cm. These dimensions are suitable for the drawing. As a check on the accuracy, compute the capacitance per meter from the sketch and from the exact expression:

$$C = \frac{2\pi\epsilon}{\cosh^{-1}\left[(a^2 + b^2 - D^2)/(2ab)\right]}$$

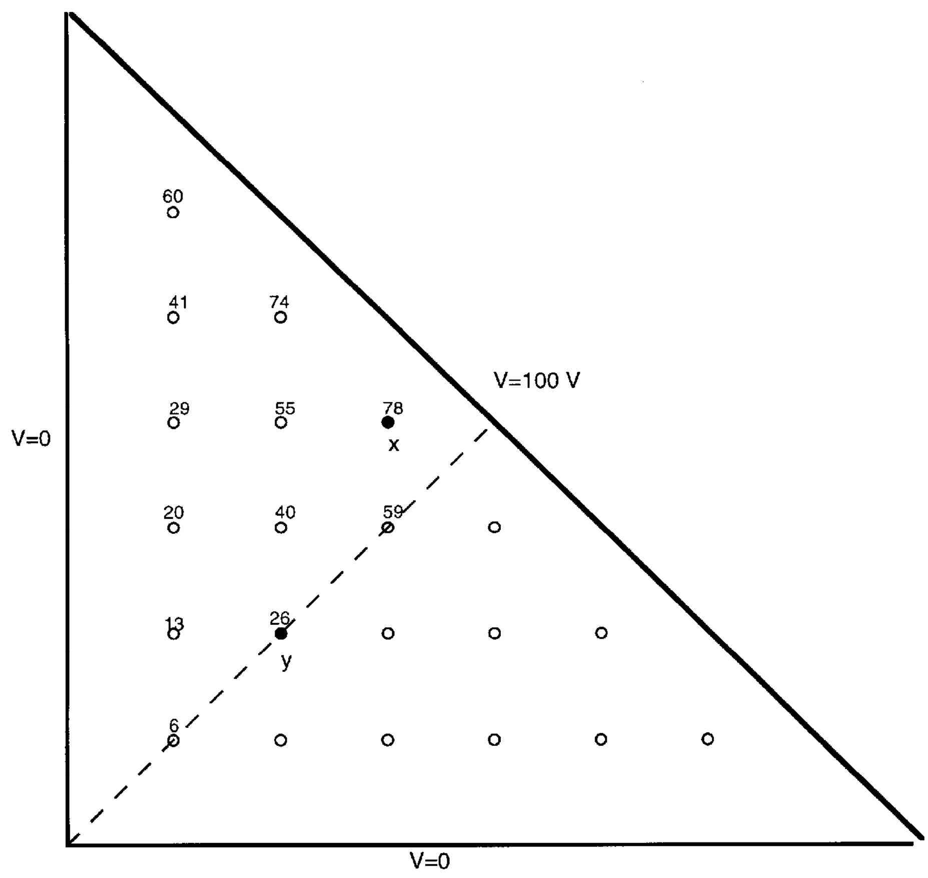
where a and b are the conductor radii and D is the axis separation.

Our attempt at the drawing is shown below. Use of the exact expression above yields a capacitance value of $C = \underline{11.5\epsilon_0 \text{ F/m}}$. Use of the drawing produces:

$$C \doteq \frac{22 \times 2}{4}\epsilon_0 = \underline{11\epsilon_0 \text{ F/m}}$$



6.7. Use the iteration method to estimate the potentials at points x and y in the triangular trough of Fig. 6.13. Work only to the nearest volt: The result is shown below. The mirror image of the values shown occur at the points on the other side of the line of symmetry (dashed line). Note that $V_x = \underline{78\,\text{V}}$ and $V_y = \underline{26\,\text{V}}$.



6.17. A two-wire transmission line consists of two parallel perfectly-conducting cylinders, each having a radius of 0.2 mm, separated by center-to-center distance of 2 mm. The medium surrounding the wires has $\epsilon_R = 3$ and $\sigma = 1.5$ mS/m. A 100-V battery is connected between the wires. Calculate:

a) the magnitude of the charge per meter length on each wire: Use

$$C = \frac{\pi\epsilon}{\cosh^{-1}(h/b)} = \frac{\pi \times 3 \times 8.85 \times 10^{-12}}{\cosh^{-1}(1/0.2)} = 3.64 \times 10^{-9} \text{ C/m}$$

Then the charge per unit length will be

$$Q = CV_0 = (3.64 \times 10^{-11})(100) = 3.64 \times 10^{-9} \text{ C/m} = \underline{3.64\,\text{nC/m}}$$

b) the battery current: Use

$$RC = \frac{\epsilon}{\sigma} \quad \Rightarrow \quad R = \frac{3 \times 8.85 \times 10^{-12}}{(1.5 \times 10^{-3})(3.64 \times 10^{-11})} = 486\,\Omega$$

Then

$$I = \frac{V_0}{R} = \frac{100}{486} = 0.206\text{ A} = \underline{206\,\text{mA}}$$

7.3. Let $V(x,y) = 4e^{2x} + f(x) - 3y^2$ in a region of free space where $\rho_v = 0$. It is known that both E_x and V are zero at the origin. Find $f(x)$ and $V(x,y)$: Since $\rho_v = 0$, we know that $\nabla^2 V = 0$, and so

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 16e^{2x} + \frac{d^2 f}{dx^2} - 6 = 0$$

Therefore

$$\frac{d^2 f}{dx^2} = -16e^{2x} + 6 \quad \Rightarrow \quad \frac{df}{dx} = -8e^{2x} + 6x + C_1$$

Now

$$E_x = \frac{\partial V}{\partial x} = 8e^{2x} + \frac{df}{dx}$$

and at the origin, this becomes

$$E_x(0) = 8 + \left.\frac{df}{dx}\right|_{x=0} = 0(\text{as given})$$

Thus $df/dx\,|_{x=0} = -8$, and so it follows that $C_1 = 0$. Integrating again, we find

$$f(x,y) = -4e^{2x} + 3x^2 + C_2$$

which at the origin becomes $f(0,0) = -4 + C_2$. However, $V(0,0) = 0 = 4 + f(0,0)$. So $f(0,0) = -4$ and $C_2 = 0$. Finally, $f(x,y) = \underline{-4e^{2x} + 3x^2}$, and $V(x,y) = 4e^{2x} - 4e^{2x} + 3x^2 - 3y^2 = \underline{3(x^2 - y^2)}$.

7.13. Coaxial conducting cylinders are located at $\rho = 0.5$ cm and $\rho = 1.2$ cm. The region between the cylinders is filled with a homogeneous perfect dielectric. If the inner cylinder is at 100V and the outer at 0V, find:

a) the location of the 20V equipotential surface: From Eq. (16) we have

$$V(\rho) = 100\frac{\ln(.012/\rho)}{\ln(.012/.005)} \text{ V}$$

We seek ρ at which $V = 20$ V, and thus we need to solve:

$$20 = 100\frac{\ln(.012/\rho)}{\ln(2.4)} \quad \Rightarrow \quad \rho = \frac{.012}{(2.4)^{0.2}} = \underline{1.01 \text{ cm}}$$

b) $E_{\rho\,max}$: We have

$$E_\rho = -\frac{\partial V}{\partial \rho} = -\frac{dV}{d\rho} = \frac{100}{\rho \ln(2.4)}$$

whose maximum value will occur at the inner cylinder, or at $\rho = .5$ cm:

$$E_{\rho\,max} = \frac{100}{.005\ln(2.4)} = 2.28 \times 10^4 \text{ V/m} = \underline{22.8 \text{ kV/m}}$$

c) ϵ_R if the charge per meter length on the inner cylinder is 20 nC/m: The capacitance per meter length is

$$C = \frac{2\pi\epsilon_0\epsilon_R}{\ln(2.4)} = \frac{Q}{V_0}$$

We solve for ϵ_R:

$$\epsilon_R = \frac{(20 \times 10^{-9})\ln(2.4)}{2\pi\epsilon_0(100)} = \underline{3.15}$$

or

$$\frac{d}{dr}\left(r^2\frac{dV}{dr}\right) = -200r^{-.4}$$

Integrate once:

$$\left(r^2\frac{dV}{dr}\right) = -\frac{200}{.6}r^{.6} + C_1 = -333.3r^{.6} + C_1$$

or

$$\frac{dV}{dr} = -333.3r^{-1.4} + \frac{C_1}{r^2} = \nabla V \text{ (in this case)} = -E_r$$

Our first boundary condition states that $r^2E_r \rightarrow 0$ when $r \rightarrow 0$ Therefore $C_1 = 0$. Integrate again to find:

$$V(r) = \frac{333.3}{.4}r^{-.4} + C_2$$

From our second boundary condition, $V \rightarrow 0$ as $r \rightarrow \infty$, we see that $C_2 = 0$. Finally,

$$V(r) = \underline{833.3r^{-.4} \text{ V}}$$

b) Now find $V(r)$ by using Gauss' Law and a line integral: Gauss' law applied to a spherical surface of radius r gives:

$$4\pi r^2 D_r = 4\pi \int_0^r \frac{200\epsilon_0}{(r')^{2.4}}(r')^2 dr = 800\pi\epsilon_0 \frac{r^{.6}}{.6}$$

Thus

$$E_r = \frac{D_r}{\epsilon_0} = \frac{800\pi\epsilon_0 r^{.6}}{.6(4\pi)\epsilon_0 r^2} = 333.3 r^{-1.4} \text{ V/m}$$

Now

$$V(r) = -\int_\infty^r 333.3(r')^{-1.4} dr' = \underline{833.3 r^{-.4} \text{ V}}$$

8.1a. Find **H** in cartesian components at $P(2,3,4)$ if there is a current filament on the z axis carrying 8 mA in the $\mathbf{a}_z$ direction:
Applying the Biot-Savart Law, we obtain

$$\mathbf{H}_a = \int_{-\infty}^{\infty} \frac{I d\mathbf{L} \times \mathbf{a}_R}{4\pi R^2} = \int_{-\infty}^{\infty} \frac{I dz\, \mathbf{a}_z \times [2\mathbf{a}_x + 3\mathbf{a}_y + (4-z)\mathbf{a}_z]}{4\pi(z^2 - 8z + 29)^{3/2}} = \int_{-\infty}^{\infty} \frac{I dz [2\mathbf{a}_y - 3\mathbf{a}_x]}{4\pi(z^2 - 8z + 29)^{3/2}}$$

Using integral tables, this evaluates as

$$\mathbf{H}_a = \frac{I}{4\pi}\left[\frac{2(2z-8)(2\mathbf{a}_y - 3\mathbf{a}_x)}{52(z^2 - 8z + 29)^{1/2}}\right]_{-\infty}^{\infty} = \frac{I}{26\pi}(2\mathbf{a}_y - 3\mathbf{a}_x)$$

Then with $I = 8$ mA, we finally obtain $\mathbf{H}_a = \underline{-294\mathbf{a}_x + 196\mathbf{a}_y \ \mu\text{A/m}}$

b. Repeat if the filament is located at $x = -1$, $y = 2$: In this case the Biot-Savart integral becomes

$$\mathbf{H}_b = \int_{-\infty}^{\infty} \frac{I dz\, \mathbf{a}_z \times [(2+1)\mathbf{a}_x + (3-2)\mathbf{a}_y + (4-z)\mathbf{a}_z]}{4\pi(z^2 - 8z + 26)^{3/2}} = \int_{-\infty}^{\infty} \frac{I dz [3\mathbf{a}_y - \mathbf{a}_x]}{4\pi(z^2 - 8z + 26)^{3/2}}$$

Evaluating as before, we obtain with $I = 8$ mA:

$$\mathbf{H}_b = \frac{I}{4\pi}\left[\frac{2(2z-8)(3\mathbf{a}_y - \mathbf{a}_x)}{40(z^2 - 8z + 26)^{1/2}}\right]_{-\infty}^{\infty} = \frac{I}{20\pi}(3\mathbf{a}_y - \mathbf{a}_x) = \underline{-127\mathbf{a}_x + 382\mathbf{a}_y \ \mu\text{A/m}}$$

c. Find **H** if both filaments are present: This will be just the sum of the results of parts a and b, or

$$\mathbf{H}_T = \mathbf{H}_a + \mathbf{H}_b = \underline{-421\mathbf{a}_x + 578\mathbf{a}_y \ \mu\text{A/m}}$$

This problem can also be done (somewhat more simply) by using the known result for **H** from an infinitely-long wire in cylindrical components, and transforming to cartesian components. The Biot-Savart method was used here for the sake of illustration.

8.23. Given the field $\mathbf{H} = 20\rho^2\, \mathbf{a}_\phi$ A/m:

a) Determine the current density **J**: This is found through the curl of **H**, which simplifies to a single term, since **H** varies only with ρ and has only a ϕ component:

$$\mathbf{J} = \nabla \times \mathbf{H} = \frac{1}{\rho}\frac{d(\rho H_\phi)}{d\rho}\, \mathbf{a}_z = \frac{1}{\rho}\frac{d}{d\rho}\left(20\rho^3\right)\, \mathbf{a}_z = \underline{60\rho\, \mathbf{a}_z \ \text{A/m}^2}$$

b) Integrate $\mathbf{J}$ over the circular surface $\rho = 1$, $0 < \phi < 2\pi$, $z = 0$, to determine the total current passing through that surface in the $\mathbf{a}_z$ direction: The integral is:

$$I = \int\int \mathbf{J} \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^1 60\rho \mathbf{a}_z \cdot \rho \, d\rho \, d\phi \mathbf{a}_z = \underline{40\pi \text{ A}}$$

c) Find the total current once more, this time by a line integral around the circular path $\rho = 1$, $0 < \phi < 2\pi$, $z = 0$:

$$I = \oint \mathbf{H} \cdot d\mathbf{L} = \int_0^{2\pi} 20\rho^2 \, \mathbf{a}_\phi \Big|_{\rho=1} \cdot (1) d\phi \mathbf{a}_\phi = \int_0^{2\pi} 20 \, d\phi = \underline{40\pi \text{ A}}$$

8.41. Assume that $\mathbf{A} = 50\rho^2 \mathbf{a}_z$ Wb/m in a certain region of free space.

a) Find $\mathbf{H}$ and $\mathbf{B}$: Use

$$\mathbf{B} = \nabla \times \mathbf{A} = -\frac{\partial A_z}{\partial \rho} \mathbf{a}_\phi = \underline{-100\rho \, \mathbf{a}_\phi \text{ Wb/m}^2}$$

Then $\mathbf{H} = \mathbf{B}/\mu_0 = \underline{-100\rho/\mu_0 \, \mathbf{a}_\phi \text{ A/m}}$.

b) Find $\mathbf{J}$: Use

$$\mathbf{J} = \nabla \times \mathbf{H} = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho H_\phi) \mathbf{a}_z = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\frac{-100\rho^2}{\mu_0} \right) \mathbf{a}_z = \underline{-\frac{200}{\mu_0} \mathbf{a}_z \text{ A/m}^2}$$

c) Use $\mathbf{J}$ to find the total current crossing the surface $0 \le \rho \le 1$, $0 \le \phi < 2\pi$, $z = 0$: The current is

$$I = \int\int \mathbf{J} \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^1 \frac{-200}{\mu_0} \mathbf{a}_z \cdot \mathbf{a}_z \, \rho \, d\rho \, d\phi = \frac{-200\pi}{\mu_0} \text{ A} = \underline{-500 \text{ kA}}$$

d) Use the value of H_ϕ at $\rho = 1$ to calculate $\oint \mathbf{H} \cdot d\mathbf{L}$ for $\rho = 1$, $z = 0$: Have

$$\oint \mathbf{H} \cdot d\mathbf{L} = I = \int_0^{2\pi} \frac{-100}{\mu_0} \mathbf{a}_\phi \cdot \mathbf{a}_\phi \, (1) d\phi = \frac{-200\pi}{\mu_0} \text{ A} = \underline{-500 \text{ kA}}$$

9.3. A point charge for which $Q = 2 \times 10^{-16}$ C and $m = 5 \times 10^{-26}$ kg is moving in the combined fields $\mathbf{E} = 100\mathbf{a}_x - 200\mathbf{a}_y + 300\mathbf{a}_z$ V/m and $\mathbf{B} = -3\mathbf{a}_x + 2\mathbf{a}_y - \mathbf{a}_z$ mT. If the charge velocity at $t = 0$ is $\mathbf{v}(0) = (2\mathbf{a}_x - 3\mathbf{a}_y - 4\mathbf{a}_z) \times 10^5$ m/s:

a) give the unit vector showing the direction in which the charge is accelerating at $t = 0$: Use $\mathbf{F}(t = 0) = q[\mathbf{E} + (\mathbf{v}(0) \times \mathbf{B})]$, where

$$\mathbf{v}(0) \times \mathbf{B} = (2\mathbf{a}_x - 3\mathbf{a}_y - 4\mathbf{a}_z)10^5 \times (-3\mathbf{a}_x + 2\mathbf{a}_y - \mathbf{a}_z)10^{-3} = 1100\mathbf{a}_x + 1400\mathbf{a}_y - 500\mathbf{a}_z$$

So the force in newtons becomes

$$\mathbf{F}(0) = (2 \times 10^{-16})[(100 + 1100)\mathbf{a}_x + (1400 - 200)\mathbf{a}_y + (300 - 500)\mathbf{a}_z] = 4 \times 10^{-14}[6\mathbf{a}_x + 6\mathbf{a}_y - \mathbf{a}_z]$$

The unit vector that gives the acceleration direction is found from the force to be

$$\mathbf{a}_F = \frac{6\mathbf{a}_x + 6\mathbf{a}_y - \mathbf{a}_z}{\sqrt{73}} = \underline{.70\mathbf{a}_x + .70\mathbf{a}_y - .12\mathbf{a}_z}$$

b) find the kinetic energy of the charge at $t = 0$:

$$\text{K.E.} = \frac{1}{2}m|\mathbf{v}(0)|^2 = \frac{1}{2}(5 \times 10^{-26}\,\text{kg})(5.39 \times 10^5\,\text{m/s})^2 = 7.25 \times 10^{-15}\,\text{J} = \underline{7.25\,\text{fJ}}$$

9.15. A solid conducting filament extends from $x = -b$ to $x = b$ along the line $y = 2$, $z = 0$. This filament carries a current of 3 A in the $\mathbf{a}_x$ direction. An infinite filament on the z axis carries 5 A in the $\mathbf{a}_z$ direction. Obtain an expression for the torque exerted on the finite conductor about an origin located at $(0, 2, 0)$: The differential force on the wire segment arising from the field from the infinite wire is

$$d\mathbf{F} = 3\,dx\,\mathbf{a}_x \times \frac{5\mu_0}{2\pi\rho}\,\mathbf{a}_\phi = -\frac{15\mu_0 \cos\phi\,dx}{2\pi\sqrt{x^2+4}}\,\mathbf{a}_z = -\frac{15\mu_0 x\,dx}{2\pi(x^2+4)}\,\mathbf{a}_z$$

So now the differential torque about the $(0, 2, 0)$ origin is

$$d\mathbf{T} = \mathbf{R}_T \times d\mathbf{F} = x\,\mathbf{a}_x \times -\frac{15\mu_0 x\,dx}{2\pi(x^2+4)}\,\mathbf{a}_z = \frac{15\mu_0 x^2\,dx}{2\pi(x^2+4)}\,\mathbf{a}_y$$

The torque is then

$$\mathbf{T} = \int_{-b}^{b} \frac{15\mu_0 x^2\,dx}{2\pi(x^2+4)}\,\mathbf{a}_y = \frac{15\mu_0}{2\pi}\,\mathbf{a}_y \left[x - 2\tan^{-1}\left(\frac{x}{2}\right)\right]_{-b}^{b}$$

$$= \underline{(6 \times 10^{-6})\left[b - 2\tan^{-1}\left(\frac{b}{2}\right)\right]\mathbf{a}_y \quad \text{N}\cdot\text{m}}$$

9.33. A toroidal core has a square cross section, 2.5 cm $< \rho <$ 3.5 cm, -0.5 cm $< z <$ 0.5 cm. The upper half of the toroid, $0 < z < 0.5$ cm, is constructed of a linear material for which $\mu_R = 10$, while the lower half, -0.5 cm $< z < 0$, has $\mu_R = 20$. An mmf of 150 A · t establishes a flux in the $\mathbf{a}_\phi$ direction. For $z > 0$, find:

a) $H_\phi(\rho)$: Ampere's circuital law gives:

$$2\pi\rho H_\phi = NI = 150 \quad \Rightarrow \quad H_\phi = \frac{150}{2\pi\rho} = \underline{23.9/\rho\ \text{A/m}}$$

b) $B_\phi(\rho)$: We use $B_\phi = \mu_R\mu_0 H_\phi = (10)(4\pi \times 10^{-7})(23.9/\rho) = \underline{3.0 \times 10^{-4}/\rho\ \text{Wb/m}^2}$.

c) $\Phi_{z>0}$: This will be

$$\Phi_{z>0} = \int\int \mathbf{B}\cdot d\mathbf{S} = \int_0^{.005}\int_{.025}^{.035} \frac{3.0 \times 10^{-4}}{\rho}\,d\rho dz = (.005)(3.0 \times 10^{-4})\ln\left(\frac{.035}{.025}\right)$$
$$= \underline{5.0 \times 10^{-7}\ \text{Wb}}$$

d) Repeat for $z < 0$: First, the magnetic field strength will be the same as in part a, since the calculation is material-independent. Thus $H_\phi = \underline{23.9/\rho\ \text{A/m}}$. Next, B_ϕ is modified only by the new permeability, which is twice the value used in part *a*: Thus $B_\phi = \underline{6.0 \times 10^{-4}/\rho\ \text{Wb/m}^2}$. Finally, since B_ϕ is twice that of part *a*, the flux will be increased by the same factor, since the area of integration for $z < 0$ is the same. Thus $\Phi_{z<0} = \underline{1.0 \times 10^{-6}\ \text{Wb}}$.

e) Find Φ_{total}: This will be the sum of the values found for $z < 0$ and $z > 0$, or $\Phi_{\text{total}} = \underline{1.5 \times 10^{-6}\ \text{Wb}}$.

Supplemental Solutions for E-text

10.1. In Fig. 10.4, let $B = 0.2\cos 120\pi t$ T, and assume that the conductor joining the two ends of the resistor is perfect. It may be assumed that the magnetic field produced by $I(t)$ is negligible. Find:

a) $V_{ab}(t)$: Since B is constant over the loop area, the flux is $\Phi = \pi(0.15)^2 B = 1.41 \times 10^{-2}\cos 120\pi t$ Wb. Now, $emf = V_{ba}(t) = -d\Phi/dt = (120\pi)(1.41 \times 10^{-2})\sin 120\pi t$. Then $V_{ab}(t) = -V_{ba}(t) = \underline{-5.33\sin 120\pi t \text{ V}}$.

b) $I(t) = V_{ba}(t)/R = 5.33\sin(120\pi t)/250 = \underline{21.3\sin(120\pi t) \text{ mA}}$

10.9. A square filamentary loop of wire is 25 cm on a side and has a resistance of 125 Ω per meter length. The loop lies in the $z = 0$ plane with its corners at $(0, 0, 0)$, $(0.25, 0, 0)$, $(0.25, 0.25, 0)$, and $(0, 0.25, 0)$ at $t = 0$. The loop is moving with velocity $v_y = 50$ m/s in the field $B_z = 8\cos(1.5 \times 10^8 t - 0.5x)$ μT. Develop a function of time which expresses the ohmic power being delivered to the loop: First, since the field does not vary with y, the loop motion in the y direction does not produce any time-varying flux, and so this motion is immaterial. We can evaluate the flux at the original loop position to obtain:

$$\Phi(t) = \int_0^{.25}\int_0^{.25} 8 \times 10^{-6}\cos(1.5 \times 10^8 t - 0.5x)\, dx\, dy$$
$$= -(4 \times 10^{-6})\left[\sin(1.5 \times 10^8 t - 0.13x) - \sin(1.5 \times 10^8 t)\right] \text{ Wb}$$

Now, $emf = V(t) = -d\Phi/dt = 6.0 \times 10^2\left[\cos(1.5 \times 10^8 t - 0.13x) - \cos(1.5 \times 10^8 t)\right]$, The total loop resistance is $R = 125(0.25 + 0.25 + 0.25 + 0.25) = 125\,\Omega$. Then the ohmic power is

$$P(t) = \frac{V^2(t)}{R} = \underline{2.9 \times 10^3\left[\cos(1.5 \times 10^8 t - 0.13x) - \cos(1.5 \times 10^8 t)\right] \text{ Watts}}$$

10.17. The electric field intensity in the region $0 < x < 5$, $0 < y < \pi/12$, $0 < z < 0.06$ m in free space is given by $\mathbf{E} = C\sin(12y)\sin(az)\cos(2 \times 10^{10} t)\,\mathbf{a}_x$ V/m. Beginning with the $\nabla \times \mathbf{E}$ relationship, use Maxwell's equations to find a numerical value for a, if it is known that a is greater than zero: In this case we find

$$\nabla \times \mathbf{E} = \frac{\partial E_x}{\partial z}\mathbf{a}_y - \frac{\partial E_z}{\partial y}\mathbf{a}_z$$
$$= C\left[a\sin(12y)\cos(az)\mathbf{a}_y - 12\cos(12y)\sin(az)\mathbf{a}_z\right]\cos(2 \times 10^{10} t) = -\frac{\partial \mathbf{B}}{\partial t}$$

Then

$$\mathbf{H} = -\frac{1}{\mu_0}\int \nabla \times \mathbf{E}\; dt + C_1$$
$$= -\frac{C}{\mu_0(2 \times 10^{10}}\left[a\sin(12y)\cos(az)\mathbf{a}_y - 12\cos(12y)\sin(az)\mathbf{a}_z\right]\sin(2 \times 10^{10} t) \text{ A/m}$$

where the integration constant, $C_1 = 0$, since there are no initial conditions. Using this result, we now find

$$\nabla \times \mathbf{H} = \left[\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right]\mathbf{a}_x = -\frac{C(144 + a^2)}{\mu_0(2 \times 10^{10})}\sin(12y)\sin(az)\sin(2 \times 10^{10} t)\,\mathbf{a}_x = \frac{\partial \mathbf{D}}{\partial t}$$

10.17. (continued) Now

$$\mathbf{E} = \frac{\mathbf{D}}{\epsilon_0} = \int \frac{1}{\epsilon_0} \nabla \times \mathbf{H}\, dt + C_2 = \frac{C(144 + a^2)}{\mu_0 \epsilon_0 (2 \times 10^{10})^2} \sin(12y) \sin(az) \cos(2 \times 10^{10} t)\, \mathbf{a}_x$$

where $C_2 = 0$. This field must be the same as the original field as stated, and so we require that

$$\frac{C(144 + a^2)}{\mu_0 \epsilon_0 (2 \times 10^{10})^2} = 1$$

Using $\mu_0 \epsilon_0 = (3 \times 10^8)^{-2}$, we find

$$a = \left[\frac{(2 \times 10^{10})^2}{(3 \times 10^8)^2} - 144 \right]^{1/2} = \underline{66}$$

11.7. The phasor magnetic field intensity for a 400-MHz uniform plane wave propagating in a certain lossless material is $(2\mathbf{a}_y - j5\mathbf{a}_z)e^{-j25x}$ A/m. Knowing that the maximum amplitude of $\mathbf{E}$ is 1500 V/m, find β, η, λ, v_p, ϵ_R, μ_R, and $\mathbf{H}(x, y, z, t)$: First, from the phasor expression, we identify $\beta = \underline{25\ \text{m}^{-1}}$ from the argument of the exponential function. Next, we evaluate $H_0 = |\mathbf{H}| = \sqrt{\mathbf{H} \cdot \mathbf{H}^*} = \sqrt{2^2 + 5^2} = \sqrt{29}$. Then $\eta = E_0/H_0 = 1500/\sqrt{29} = \underline{278.5\ \Omega}$. Then $\lambda = 2\pi/\beta = 2\pi/25 = .25$ m $= \underline{25\text{ cm}}$. Next,

$$v_p = \frac{\omega}{\beta} = \frac{2\pi \times 400 \times 10^6}{25} = \underline{1.01 \times 10^8\ \text{m/s}}$$

Now we note that

$$\eta = 278.5 = 377 \sqrt{\frac{\mu_R}{\epsilon_R}} \quad \Rightarrow \quad \frac{\mu_R}{\epsilon_R} = 0.546$$

And

$$v_p = 1.01 \times 10^8 = \frac{c}{\sqrt{\mu_R \epsilon_R}} \quad \Rightarrow \quad \mu_R \epsilon_R = 8.79$$

We solve the above two equations simultaneously to find $\epsilon_R = \underline{4.01}$ and $\mu_R = \underline{2.19}$. Finally,

$$\begin{aligned} \mathbf{H}(x, y, z, t) &= Re\left\{(2\mathbf{a}_y - j5\mathbf{a}_z)e^{-j25x} e^{j\omega t}\right\} \\ &= \underline{2\cos(2\pi \times 400 \times 10^6 t - 25x)\mathbf{a}_y + 5\sin(2\pi \times 400 \times 10^6 t - 25x)\mathbf{a}_z\ \text{A/m}} \end{aligned}$$

11.15. A 10 GHz radar signal may be represented as a uniform plane wave in a sufficiently small region. Calculate the wavelength in centimeters and the attenuation in nepers per meter if the wave is propagating in a non-magnetic material for which

a) $\epsilon'_R = 1$ and $\epsilon''_R = 0$: In a non-magnetic material, we would have:

$$\alpha = \omega \sqrt{\frac{\mu_0 \epsilon_0 \epsilon'_R}{2}} \left[\sqrt{1 + \left(\frac{\epsilon''_R}{\epsilon'_R} \right)^2} - 1 \right]^{1/2}$$

and

$$\beta = \omega \sqrt{\frac{\mu_0 \epsilon_0 \epsilon'_R}{2}} \left[\sqrt{1 + \left(\frac{\epsilon''_R}{\epsilon'_R} \right)^2} + 1 \right]^{1/2}$$

With the given values of ϵ'_R and ϵ''_R, it is clear that $\beta = \omega\sqrt{\mu_0 \epsilon_0} = \omega/c$, and so $\lambda = 2\pi/\beta = 2\pi c/\omega = 3 \times 10^{10}/10^{10} = \underline{3\text{ cm}}$. It is also clear that $\alpha = \underline{0}$.

11.15. (continued)

b) $\epsilon'_R = 1.04$ and $\epsilon''_R = 9.00 \times 10^{-4}$: In this case $\epsilon''_R/\epsilon'_R << 1$, and so $\beta \doteq \omega\sqrt{\epsilon'_R}/c = 2.13\ \text{cm}^{-1}$. Thus $\lambda = 2\pi/\beta = \underline{2.95\ \text{cm}}$. Then

$$\alpha \doteq \frac{\omega\epsilon''}{2}\sqrt{\frac{\mu}{\epsilon'}} = \frac{\omega\epsilon''_R}{2}\frac{\sqrt{\mu_0\epsilon_0}}{\sqrt{\epsilon'_R}} = \frac{\omega}{2c}\frac{\epsilon''_R}{\sqrt{\epsilon'_R}} = \frac{2\pi\times 10^{10}}{2\times 3\times 10^8}\frac{(9.00\times 10^{-4})}{\sqrt{1.04}}$$
$$= \underline{9.24\times 10^{-2}\ \text{Np/m}}$$

c) $\epsilon'_R = 2.5$ and $\epsilon''_R = 7.2$: Using the above formulas, we obtain

$$\beta = \frac{2\pi\times 10^{10}\sqrt{2.5}}{(3\times 10^{10})\sqrt{2}}\left[\sqrt{1+\left(\frac{7.2}{2.5}\right)^2}+1\right]^{1/2} = 4.71\ \text{cm}^{-1}$$

and so $\lambda = 2\pi/\beta = \underline{1.33\ \text{cm}}$. Then

$$\alpha = \frac{2\pi\times 10^{10}\sqrt{2.5}}{(3\times 10^{8})\sqrt{2}}\left[\sqrt{1+\left(\frac{7.2}{2.5}\right)^2}-1\right]^{1/2} = \underline{335\ \text{Np/m}}$$

11.29. Consider a left-circularly polarized wave in free space that propagates in the forward z direction. The electric field is given by the appropriate form of Eq. (80).

a) Determine the magnetic field phasor, $\mathbf{H}_s$:
We begin, using (80), with $\mathbf{E}_s = E_0(\mathbf{a}_x + j\mathbf{a}_y)e^{-j\beta z}$. We find the two components of $\mathbf{H}_s$ separately, using the two components of $\mathbf{E}_s$. Specifically, the x component of $\mathbf{E}_s$ is associated with a y component of $\mathbf{H}_s$, and the y component of $\mathbf{E}_s$ is associated with a negative x component of $\mathbf{H}_s$. The result is

$$\mathbf{H}_s = \underline{\frac{E_0}{\eta_0}(\mathbf{a}_y - j\mathbf{a}_x)\,e^{-j\beta z}}$$

b) Determine an expression for the average power density in the wave in W/m^2 by direct application of Eq. (57): We have

$$\mathcal{P}_{z,avg} = \frac{1}{2}Re(\mathbf{E}_s\times\mathbf{H}_s^*) = \frac{1}{2}Re\left(E_0(\mathbf{a}_x + j\mathbf{a}_y)e^{-j\beta z}\times\frac{E_0}{\eta_0}(\mathbf{a}_y - j\mathbf{a}_x)e^{+j\beta z}\right)$$
$$= \underline{\frac{E_0^2}{\eta_0}\mathbf{a}_z\ \text{W/m}^2}\quad \text{(assuming } E_0 \text{ is real)}$$

12.3. A uniform plane wave in region 1 is normally incident on the planar boundary separating regions 1 and 2. If $\epsilon''_1 = \epsilon''_2 = 0$, while $\epsilon'_{R1} = \mu^3_{R1}$, and $\epsilon'_{R2} = \mu^3_{R2}$, find the ratio $\epsilon'_{R2}/\epsilon'_{R1}$ if 20% of the energy in the incident wave is reflected at the boundary. There are two possible answers: We begin with the reflection coefficient:

$$\Gamma = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1} = \frac{\sqrt{\mu_{R2}/\epsilon'_{R2}} - \sqrt{\mu_{R1}/\epsilon'_{R1}}}{\sqrt{\mu_{R2}/\epsilon'_{R2}} + \sqrt{\mu_{R1}/\epsilon'_{R1}}} = \frac{\sqrt{\mu_{R2}/\mu^3_{R2}} - \sqrt{\mu_{R1}/\mu^3_{R1}}}{\sqrt{\mu_{R2}/\mu^3_{R2}} + \sqrt{\mu_{R1}/\mu^3_{R1}}} = \frac{\mu_{R1} - \mu_{R2}}{\mu_{R1} + \mu_{R2}}$$

But we are given that $|\Gamma| = 0.20$, so that $\Gamma = \pm\sqrt{.20}$, or

$$\frac{\mu_{R2}}{\mu_{R1}} = \frac{1\mp\sqrt{.20}}{1\pm\sqrt{.20}} = 0.38 \text{ or } 2.62 \;\Rightarrow\; \frac{\epsilon'_{R2}}{\epsilon'_{R1}} = \left(\frac{\mu_{R2}}{\mu_{R1}}\right)^3 = \underline{0.05 \text{ or } 18.0}$$

12.11. A 150 MHz uniform plane wave in normally-incident from air onto a material whose intrinsic impedance is unknown. Measurements yield a standing wave ratio of 3 and the appearance of an electric field minimum at 0.3 wavelengths in front of the interface. Determine the impedance of the unknown material: First, the field minimum is used to find the phase of the reflection coefficient, where

$$z_{min} = -\frac{1}{2\beta}(\phi + \pi) = -0.3\lambda \quad \Rightarrow \quad \phi = 0.2\pi$$

where $\beta = 2\pi/\lambda$ has been used. Next,

$$|\Gamma| = \frac{s-1}{s+1} = \frac{3-1}{3+1} = \frac{1}{2}$$

So we now have

$$\Gamma = 0.5e^{j0.2\pi} = \frac{\eta_u - \eta_0}{\eta_u + \eta_0}$$

We solve for η_u to find

$$\eta_u = \eta_0(1.70 + j1.33) = \underline{641 + j501\ \Omega}$$

12.21. A right-circularly polarized plane wave in air is incident at Brewster's angle onto a semi-infinite slab of plexiglas ($\epsilon'_R = 3.45$, $\epsilon''_R = 0$, $\mu = \mu_0$).

a) Determine the fractions of the incident power that are reflected and transmitted: In plexiglas, Brewster's angle is $\theta_B = \theta_1 = \tan^{-1}(\epsilon'_{R2}/\epsilon'_{R1}) = \tan^{-1}(\sqrt{3.45}) = 61.7°$. Then the angle of refraction is $\theta_2 = 90° - \theta_B$ (see Example 12.9), or $\theta_2 = 28.3°$. With incidence at Brewster's angle, all p-polarized power will be transmitted — only s-polarized power will be reflected. This is found through

$$\Gamma_s = \frac{\eta_{2s} - \eta_{1s}}{\eta_{2s} + \eta_{1s}} = \frac{.614\eta_0 - 2.11\eta_0}{.614\eta_0 + 2.11\eta_0} = -0.549$$

where $\eta_{1s} = \eta_1 \sec\theta_1 = \eta_0 \sec(61.7°) = 2.11\eta_0$, and $\eta_{2s} = \eta_2 \sec\theta_2 = (\eta_0/\sqrt{3.45})\sec(28.3°) = 0.614\eta_0$. Now, the reflected power fraction is $|\Gamma|^2 = (-.549)^2 = .302$. Since the wave is circularly-polarized, the s-polarized component represents one-half the total incident wave power, and so the fraction of the *total* power that is reflected is $.302/2 = 0.15$, or 15%. The fraction of the incident power that is transmitted is then the remainder, or 85%.

b) Describe the polarizations of the reflected and transmitted waves: Since all the p-polarized component is transmitted, the reflected wave will be entirely s-polarized (linear). The transmitted wave, while having all the incident p-polarized power, will have a reduced s-component, and so this wave will be right-elliptically polarized.

13.1. The parameters of a certain transmission line operating at 6×10^8 rad/s are $L = 0.4\ \mu$H/m, $C = 40$ pF/m, $G = 80$ mS/m, and $R = 20\ \Omega/m$.

a) Find γ, α, β, λ, and Z_0: We use

$$\gamma = \sqrt{ZY} = \sqrt{(R + j\omega L)(G + j\omega C)}$$
$$= \sqrt{[20 + j(6\times 10^8)(0.4\times 10^{-6})][80\times 10^{-3} + j(6\times 10^8)(40\times 10^{-12})]}$$
$$= \underline{2.8 + j3.5\ \text{m}^{-1}} = \alpha + j\beta$$

Therefore, $\alpha = \underline{2.8\ \text{Np/m}}$, $\beta = \underline{3.5\ \text{rad/m}}$, and $\lambda = 2\pi/\beta = \underline{1.8\ \text{m}}$. Finally,

$$Z_0 = \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}} = \sqrt{\frac{20 + j2.4\times 10^2}{80\times 10^{-3} + j2.4\times 10^{-2}}} = \underline{44 + j30\ \Omega}$$

13.1. (continued)

b) If a voltage wave travels 20 m down the line, by what percentage is its amplitude reduced, and by how many degress is its phase shifted? First,

$$\frac{V_{20}}{V_0} = e^{-\alpha L} = e^{-(2.8)(20)} = 4.8 \times 10^{-25} \text{ or } \underline{4.8 \times 10^{-23} \text{ percent!}}$$

Then the phase shift is given by βL, which in degrees becomes

$$\phi = \beta L \left(\frac{360}{2\pi}\right) = (3.5)(20)\left(\frac{360}{2\pi}\right) = \underline{4.0 \times 10^3 \text{ degrees}}$$

13.21. A lossless line having an air dielectric has a characteristic impedance of 400 Ω. The line is operating at 200 MHz and $Z_{in} = 200 - j200$ Ω. Use analytic methods or the Smith chart (or both) to find: (a) s; (b) Z_L if the line is 1 m long; (c) the distance from the load to the nearest voltage maximum: I will use the analytic approach. Using normalized impedances, Eq. (13) becomes

$$z_{in} = \frac{Z_{in}}{Z_0} = \left[\frac{z_L \cos(\beta L) + j\sin(\beta L)}{\cos(\beta L) + jz_L \sin(\beta L)}\right] = \left[\frac{z_L + j\tan(\beta L)}{1 + jz_L \tan(\beta L)}\right]$$

Solve for z_L:

$$z_L = \left[\frac{z_{in} - j\tan(\beta L)}{1 - jz_{in}\tan(\beta L)}\right]$$

where, with $\lambda = c/f = 3 \times 10^8/2 \times 10^8 = 1.50$ m, we find $\beta L = (2\pi)(1)/(1.50) = 4.19$, and so $\tan(\beta L) = 1.73$. Also, $z_{in} = (200 - j200)/400 = 0.5 - j0.5$. So

$$z_L = \frac{0.5 - j0.5 - j1.73}{1 - j(0.5 - j0.5)(1.73)} = 2.61 + j0.173$$

Finally, $Z_L = z_L(400) = \underline{1.05 \times 10^3 + j69.2\ \Omega}$. Next

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{6.45 \times 10^2 + j69.2}{1.45 \times 10^3 + j69.2} = .448 + j2.64 \times 10^{-2} = .449\angle 5.9 \times 10^{-2} \text{ rad}$$

Now

$$s = \frac{1 + |\Gamma|}{1 - |\Gamma|} = \frac{1 + .449}{1 - .449} = \underline{2.63}$$

Finally

$$z_{max} = -\frac{\phi}{2\beta} = -\frac{\lambda\phi}{4\pi} = -\frac{(5.9 \times 10^{-2})(1.50)}{4\pi} = -7.0 \times 10^{-3} \text{ m} = \underline{-7.0 \text{ mm}}$$

13.37. In the transmission line of Fig. 13.17, $R_L = Z_0 = 50$ Ω. Determine and plot the voltage at the load resistor and the current in the battery as functions of time by constructing appropriate voltage and current reflection diagrams: Referring to the figure, closing the switch launches a voltage wave whose value is given by Eq. (50):

$$V_1^+ = \frac{V_0 Z_0}{R_g + Z_0} = \frac{50}{75}V_0 = \frac{2}{3}V_0$$

13.37. (continued) We note that $\Gamma_L = 0$, since the load impedance is matched to that of the line. So the voltage wave traverses the line and does not reflect. The voltage reflection diagram would be that shown in Fig. 13.18a, except that no waves are present after time $t = l/v$. Likewise, the current reflection diagram is that of Fig. 13.19a, except, again, no waves exist after $t = l/v$. The voltage at the load will be just $V_1^+ = (2/3)V_0$ for times beyond l/v. The current through the battery is found through

$$I_1^+ = \frac{V_1^+}{Z_0} = \frac{V_0}{75} \text{ A}$$

This current initiates at $t = 0$, and continues indefinitely.

14.1. A parallel-plate waveguide is known to have a cutoff wavelength for the $m = 1$ TE and TM modes of $\lambda_{c1} = 0.4$ cm. The guide is operated at wavelength $\lambda = 1$ mm. How many modes propagate? The cutoff wavelength for mode m is $\lambda_{cm} = 2nd/m$, where n is the refractive index of the guide interior. For the first mode, we are given

$$\lambda_{c1} = \frac{2nd}{1} = 0.4 \text{ cm} \quad \Rightarrow \quad d = \frac{0.4}{2n} = \frac{0.2}{n} \text{ cm}$$

Now, for mode m to propagate, we require

$$\lambda \le \frac{2nd}{m} = \frac{0.4}{m} \quad \Rightarrow \quad m \le \frac{0.4}{\lambda} = \frac{0.4}{0.1} = 4$$

So, accounting for 2 modes (TE and TM) for each value of m, we will have a total of $\underline{8}$ modes.

14.9. A rectangular waveguide has dimensions $a = 6$ cm and $b = 4$ cm.

a) Over what range of frequencies will the guide operate single mode? The cutoff frequency for mode mp is, using Eq. (54):

$$f_{c,mn} = \frac{c}{2n}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{p}{b}\right)^2}$$

where n is the refractive index of the guide interior. We require that the frequency lie between the cutoff frequencies of the TE_{10} and TE_{01} modes. These will be:

$$f_{c10} = \frac{c}{2na} = \frac{3 \times 10^8}{2n(.06)} = \frac{2.5 \times 10^9}{n}$$

$$f_{c01} = \frac{c}{2nb} = \frac{3 \times 10^8}{2n(.04)} = \frac{3.75 \times 10^9}{n}$$

Thus, the range of frequencies over which single mode operation will occur is

$$\underline{\frac{2.5}{n} \text{ GHz} < f < \frac{3.75}{n} \text{ GHz}}$$

b) Over what frequency range will the guide support *both* TE_{10} and TE_{01} modes and no others? We note first that f must be greater than f_{c01} to support both modes, but must be less than the cutoff frequency for the next higher order mode. This will be f_{c11}, given by

$$f_{c11} = \frac{c}{2n}\sqrt{\left(\frac{1}{.06}\right)^2 + \left(\frac{1}{.04}\right)^2} = \frac{30c}{2n} = \frac{4.5 \times 10^9}{n}$$

The allowed frequency range is then

$$\underline{\frac{3.75}{n} \text{ GHz} < f < \frac{4.5}{n} \text{ GHz}}$$

14.25. A dipole antenna in free space has a linear current distribution. If the length is 0.02λ, what value of I_0 is required to:

a) provide a radiation-field amplitude of 100 mV/m at a distance of one mile, at $\theta = 90°$: With a linear current distribution, the peak current, I_0, occurs at the center of the dipole; current decreases linearly to zero at the two ends. The average current is thus $I_0/2$, and we use Eq. (84) to write:

$$|E_\theta| = \frac{I_0 d \eta_0}{4\lambda r} \sin(90°) = \frac{I_0(0.02)(120\pi)}{(4)(5280)(12)(0.0254)} = 0.1 \quad \Rightarrow \quad I_0 = \underline{85.4 \text{ A}}$$

b) radiate a total power of 1 watt? We use

$$\mathcal{P}_{avg} = \left(\frac{1}{4}\right)\left(\frac{1}{2} I_0^2 R_{rad}\right)$$

where the radiation resistance is given by Eq. (86), and where the factor of 1/4 arises from the average current of $I_0/2$: We obtain:

$$\mathcal{P}_{avg} = 10\pi^2 I_0^2 (0.02)^2 = 1 \quad \Rightarrow \quad I_0 = \underline{5.03 \text{ A}}$$

APPENDIX A

VECTOR ANALYSIS

A.1 GENERAL CURVILINEAR COORDINATES

Let us consider a general orthogonal coordinate system in which a point is located by the intersection of three mutually perpendicular surfaces (of unspecified form or shape),

$$u = \text{constant}$$
$$v = \text{constant}$$
$$w = \text{constant}$$

where u, v, and w are the variables of the coordinate system. If each variable is increased by a differential amount and three more mutually perpendicular surfaces are drawn corresponding to these new values, a differential volume is formed which is closely a rectangular parallelepiped. Since u, v, and w need not be measures of length, such as, for example, the angle variables of the cylindrical and spherical coordinate systems, each must be multiplied by a general function of u, v, and w in order to obtain the differential sides of the parallelepiped. Thus we define the scale factors h_1, h_2, and h_3 each as a function of the three variables u, v, and w and write the lengths of the sides of the differential volume as

$$dL_1 = h_1 du$$
$$dL_2 = h_2 dv$$
$$dL_3 = h_3 dw$$

In the three coordinate systems discussed in Chap. 1, it is apparent that the variables and scale factors are

$$\begin{array}{llll} \text{Cartesian:} & u = x & v = y & w = z \\ & h_1 = 1 & h_2 = 1 & h_3 = 1 \\ \text{Cylindrical:} & u = \rho & v = \phi & w = z \\ & h_1 = 1 & h_2 = \rho & h_3 = 1 \\ \text{Spherical:} & u = r & v = \theta & w = \phi \\ & h_1 = 1 & h_2 = r & h_3 = r\sin\theta \end{array} \tag{A.1}$$

The choice of u, v, and w above has been made so that $\mathbf{a}_u \times \mathbf{a}_v = \mathbf{a}_w$ in all cases. More involved expressions for h_1, h_2, and h_3 are to be expected in other less familiar coordinate systems.[1]

A.2 DIVERGENCE, GRADIENT, AND CURL IN GENERAL CURVILINEAR COORDINATES

If the method used to develop divergence in Secs. 3.4 and 3.5 is applied to the general curvilinear coordinate system, the flux of the vector $\mathbf{D}$ passing through the surface of the parallelepiped whose unit normal is $\mathbf{a}_u$ is

$$D_{u0} dL_2 dL_3 + \frac{1}{2}\frac{\partial}{\partial u}(D_u dL_2 dL_3) du$$

or

$$D_{u0} h_2 h_3 dv\, dw + \frac{1}{2}\frac{\partial}{\partial u}(D_u h_2 h_3 dv\, dw) du$$

and for the opposite face it is

$$-D_{u0} h_2 h_3 dv\, dw + \frac{1}{2}\frac{\partial}{\partial u}(D_u h_2 h_3 dv\, dw) du$$

giving a total for these two faces of

$$\frac{\partial}{\partial u}(D_u h_2 h_3 dv\, dw) du$$

Since u, v, and w are independent variables, this last expression may be written as

[1] The variables and scale factors are given for nine orthogonal coordinate systems on pp. 50–59 in J. A. Stratton, "Electromagnetic Theory," McGraw-Hill Book Company, New York, 1941. Each system is also described briefly.

$$\frac{\partial}{\partial u}(h_2 h_3 D_u) du\, dv\, dw$$

and the other two corresponding expressions obtained by a simple permutation of the subscripts and of u, v, and w. Thus the total flux leaving the differential volume is

$$\left[\frac{\partial}{\partial u}(h_2 h_3 D_u) + \frac{\partial}{\partial v}(h_3 h_1 D_v) + \frac{\partial}{\partial w}(h_1 h_2 D_w)\right] du\, dv\, dw$$

and the divergence of **D** is found by dividing by the differential volume

$$\nabla \cdot \mathbf{D} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u}(h_2 h_3 D_u) + \frac{\partial}{\partial v}(h_3 h_1 D_v) + \frac{\partial}{\partial w}(h_1 h_2 D_w)\right] \tag{A.2}$$

The components of the gradient of a scalar V may be obtained (following the methods of Sec. 4.6) by expressing the total differential of V,

$$dV = \frac{\partial V}{\partial u} du + \frac{\partial V}{\partial v} dv + \frac{\partial V}{\partial w} dw$$

in terms of the component differential lengths, $h_1 du$, $h_2 dv$, and $h_3 dw$,

$$dV = \frac{1}{h_1}\frac{\partial V}{\partial u} h_1 du + \frac{1}{h_2}\frac{\partial V}{\partial v} h_2 dv + \frac{1}{h_3}\frac{\partial V}{\partial w} h_3 dw$$

Then, since

$$d\mathbf{L} = h_1 du \mathbf{a}_u + h_2 dv \mathbf{a}_v + h_3 dw \mathbf{a}_w \qquad \text{and} \qquad dV = \nabla V \cdot d\mathbf{L}$$

we see that

$$\nabla V = \frac{1}{h_1}\frac{\partial V}{\partial u}\mathbf{a}_u + \frac{1}{h_2}\frac{\partial V}{\partial v}\mathbf{a}_v + \frac{1}{h_3}\frac{\partial V}{\partial w}\mathbf{a}_w \tag{A.3}$$

The components of the curl of a vector **H** are obtained by considering a differential path first in a $u =$ constant surface and finding the circulation of **H** about that path, as discussed for cartesian coordinates in Sec. 8.3. The contribution along the segment in the $\mathbf{a}_v$ direction is

$$H_{v0} h_2 dv - \frac{1}{2}\frac{\partial}{\partial w}(H_v h_2 dv) dw$$

and that from the oppositely directed segment is

$$-H_{v0} h_2 dv - \frac{1}{2}\frac{\partial}{\partial w}(H_v h_2 dv) dw$$

The sum of these two parts is

$$-\frac{\partial}{\partial w}(H_v h_2 dv) dw$$

or

$$-\frac{\partial}{\partial w}(h_2 H_v) dv\, dw$$

and the sum of the contributions from the other two sides of the path is

$$\frac{\partial}{\partial v}(h_3 H_w) dv\, dw$$

Adding these two terms and dividing the sum by the enclosed area, $h_2 h_3 dv\, dw$, we see that the $\mathbf{a}_u$ component of curl $\mathbf{H}$ is

$$(\nabla \times \mathbf{H})_u = \frac{1}{h_2 h_3}\left[\frac{\partial}{\partial v}(h_3 H_w) - \frac{\partial}{\partial w}(h_2 H_v)\right]$$

and the other two components may be obtained by cyclic permutation. The result is expressible as a determinant,

$$\nabla \times \mathbf{H} = \begin{vmatrix} \dfrac{\mathbf{a}_u}{h_2 h_3} & \dfrac{\mathbf{a}_v}{h_3 h_1} & \dfrac{\mathbf{a}_w}{h_1 h_2} \\ \dfrac{\partial}{\partial u} & \dfrac{\partial}{\partial v} & \dfrac{\partial}{\partial w} \\ h_1 H_u & h_2 H_v & h_3 H_w \end{vmatrix} \tag{A.4}$$

The Laplacian of a scalar is found by using (2) and (3):

$$\nabla^2 V = \nabla \cdot \nabla V = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u}\left(\frac{h_2 h_3}{h_1}\frac{\partial V}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{h_3 h_1}{h_2}\frac{\partial V}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{h_1 h_2}{h_3}\frac{\partial V}{\partial w}\right)\right] \tag{A.5}$$

Equations (2) to (5) may be used to find the divergence, gradient, curl, and Laplacian in any orthogonal coordinate system for which h_1, h_2, and h_3 are known.

Expressions for $\nabla \cdot \mathbf{D}$, ∇V, $\nabla \times \mathbf{H}$, and $\nabla^2 V$ are given in cartesian, circular cylindrical, and spherical coordinate systems inside the back cover.

A.3 VECTOR IDENTITIES

The vector identities listed below may be proved by expansion in cartesian (or general curvilinear) coordinates. The first two identities involve the scalar and vector triple products, the next three are concerned with operations on sums, the following three apply to operations when the argument is multiplied by a scalar function, the next three apply to operations on scalar or vector products, and the last four concern the second-order operations.

$$(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} \equiv (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{A} \equiv (\mathbf{C} \times \mathbf{A}) \cdot \mathbf{B} \tag{A.6}$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \equiv (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C} \tag{A.7}$$

$$\nabla \cdot (\mathbf{A} + \mathbf{B}) \equiv \nabla \cdot \mathbf{A} + \nabla \cdot \mathbf{B} \tag{A.8}$$

$$\nabla(V + W) \equiv \nabla V + \nabla W \tag{A.9}$$

$$\nabla \times (\mathbf{A} + \mathbf{B}) \equiv \nabla \times \mathbf{A} + \nabla \times \mathbf{B} \tag{A.10}$$

$$\nabla \cdot (V\mathbf{A}) \equiv \mathbf{A} \cdot \nabla V + V\nabla \cdot \mathbf{A} \tag{A.11}$$

$$\nabla(VW) \equiv V\nabla W + W\nabla V \tag{A.12}$$

$$\nabla \times (V\mathbf{A}) \equiv \nabla V \times \mathbf{A} + V\nabla \times \mathbf{A} \tag{A.13}$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) \equiv \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B} \tag{A.14}$$

$$\nabla(\mathbf{A} \cdot \mathbf{B}) \equiv (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} + \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) \tag{A.15}$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) \equiv \mathbf{A}\nabla \cdot \mathbf{B} - \mathbf{B}\nabla \cdot \mathbf{A} + (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} \tag{A.16}$$

$$\nabla \cdot \nabla V \equiv \nabla^2 V \tag{A.17}$$

$$\nabla \cdot \nabla \times \mathbf{A} \equiv 0 \tag{A.18}$$

$$\nabla \times \nabla V \equiv 0 \tag{A.19}$$

$$\nabla \times \nabla \times \mathbf{A} \equiv \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \tag{A.20}$$

APPENDIX B

UNITS

We shall describe first the International System (abbreviated SI, for Système International d'Unités), which is used in this book and is now standard in electrical engineering and much of physics. It has also been officially adopted as the international system of units by many countries, including the United States.[1]

The fundamental unit of length is the meter, which was defined in the latter part of the 19th century as the distance between two marks on a certain platinum-iridium bar. The definition was improved in 1960 by relating the meter to the wavelength of the radiation emitted by the rare gas isotope krypton 86 under certain specified conditions. This so-called krypton meter was accurate to four parts per billion, a value leading to negligible uncertainties in constructing skyscrapers or building highways, but capable of causing an error greater than one meter in determining the distance to the moon. The meter was redefined in 1983 in terms of the velocity of light. At that time the velocity of light was specified to be an auxiliary constant with an *exact* value of 299 792 458 meters per second. As a result, the latest definition of the meter is the distance light travels in a vacuum

[1] The International System of Units was adopted by the Eleventh General Conference on Weights and Measures in Paris in 1960, and it was officially adopted for scientific usage by the National Bureau of Standards in 1964. It is a metric system which interestingly enough is the only system which has ever received specific sanction from Congress. This occurred first in 1966 and then again in 1975 with the Metric Conversion Act, which provides for ''voluntary conversion'' to the metric system. No specific time was specified, however, and we can assume that it will still be a few years before the bathroom scale reads mass in kilograms and Miss America is a 90–60–90.

in 1/299 792 458 of a second. If greater accuracy is achieved in measuring c, that value will remain 299 792 458 m/s, but the length of the meter will change.

It is evident that our definition of the meter is expressed in terms of the "second," the fundamental unit of time. The second is defined as 9 192 631 770 periods of the transition frequency between the hyperfine levels $F = 4$, $m_F = 0$, and $F = 3$, $m_F = 0$ of the ground state $^2s_{1/2}$ of the atom of cesium 133, unperturbed by external fields. This definition of the second, complex though it may be, permits time to be measured with an accuracy better than one part in 10^{13}.

The standard mass of one kilogram is defined as the mass of an international standard in the form of a platinum-iridium cylinder at the International Bureau of Weights and Measures at Sèvres, France.

The unit of temperature is the kelvin, defined by placing the triple-point temperature of water at 273.16 kelvins.

A fifth unit is the candela, defined as the luminous intensity of an omnidirectional radiator at the freezing temperature of platinum (2042 K) having an area of 1/600 000 square meter and under a pressure of 101 325 newtons per square meter.

The last of the fundamental units is the ampere. Before explicitly defining the ampere, we must first define the newton. It is defined in terms of the other fundamental units from Newton's third law as the force required to produce an acceleration of one meter per second per second on a one-kilogram mass. We now may define the ampere as that constant current, flowing in opposite directions in two straight parallel conductors of infinite length and negligible cross section, separated one meter in vacuum, that produces a repulsive force of 2×10^{-7} newton per meter length between the two conductors. The force between the two parallel conductors is known to be

$$F = \mu_0 \frac{I^2}{2\pi d}$$

and thus

$$2 \times 10^{-7} = \mu_0 \frac{1}{2\pi}$$

or

$$\mu_0 = 4\pi 10^{-7} \qquad (\mathrm{kg \cdot m/A^2 \cdot s^2, \text{ or } H/m})$$

We thus find that our definition of the ampere has been formulated in such a way as to assign an exact simple numerical value to the permeability of free space.

Returning to the International System, the units in which the other electric and magnetic quantities are measured are given in the body of the text at the time each quantity is defined, and all of them can be related to the basic units already defined. For example, our work with the plane wave in Chap. 11 shows that the velocity with which an electromagnetic wave propagates in free space is

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

and thus

$$\epsilon_0 = \frac{1}{\mu_0 c^2} = \frac{1}{4\pi 10^{-7} c^2} = 8.854\,187\,817 \times 10^{-12}\ \text{F/m}$$

It is evident that the numerical value of ϵ_0 depends upon the defined value of the velocity of light in vacuum, 299 792 458 m/s.

The units are also given in Table B.1 for easy reference. They are listed in the same order that they are defined in the text.

Finally, other systems of units have been used in electricity and magnetism. In the electrostatic system of units (esu), Coulomb's law is written for free space,

$$F = \frac{Q_1 Q_2}{R^2} \qquad \text{(esu)}$$

The permittivity of free space is assigned the value of unity. The gram and centimeter are the fundamental units of mass and distance, and the esu system is therefore a cgs system. Units bearing the prefix stat- belong to the electrostatic system of units.

In a similar manner, the electromagnetic system of units (emu) is based on Coulomb's law for magnetic poles, and the permeability of free space is unity. The prefix ab- identifies emu units. When electric quantities are expressed in esu units, magnetic quantities in emu units, and both appear in the same equation (such as Maxwell's curl equations), the velocity of light appears explicitly. This follows from noting that in esu $\epsilon_0 = 1$, but $\mu_0 \epsilon_0 = 1/c^2$, and therefore $\mu_0 = 1/c^2$, and in emu $\mu_0 = 1$, and hence $\epsilon_0 = 1/c^2$. Thus, in this intermixed system known as the gaussian system of units,

$$\nabla \times \mathbf{H} = 4\pi \mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} \qquad \text{(gaussian)}$$

Other systems include the factor 4π explicitly in Coulomb's law, and it then does not appear in Maxwell's equations. When this is done, the system is said to be rationalized. Hence the gaussian system is an unrationalized cgs system (when rationalized it is known as the Heaviside-Lorentz system), and the International System we have used throughout this book is a rationalized mks system.

Table B.2 gives the conversion factors between the more important units of the International System (or rationalized mks system) and the gaussian system, and several other assorted units.

Table B.3 lists the prefixes used with any of the SI units, their abbreviations, and the power of ten each represents. Those checked are widely used. Both the prefixes and their abbreviations are written without hyphens, and therefore 10^{-6} F = 1 microfarad = 1 μF = 1000 nanofarads = 1000 nF, and so forth.

TABLE B.1
Names and units of the electric and magnetic quantities in the International System (in the order they appear in the text)

Symbol	Name	Unit	Abbreviation
v	Velocity	meter/second	m/s
F	Force	newton	N
Q	Charge	coulomb	C
r, R	Distance	meter	m
ϵ_0, ϵ	Permittivity	farad/meter	F/m
E	Electric field intensity	volt/meter	V/m
ρ_v	Volume charge density	coulomb/meter^3	C/m^3
v	Volume	meter^3	m^3
ρ_L	Linear charge density	coulomb/meter	C/m
ρ_S	Surface charge density	coulomb/meter^2	C/m^2
Ψ	Electric flux	coulomb	C
D	Electric flux density	coulomb/meter^2	C/m^2
S	Area	meter^2	m^2
W	Work, energy	joule	J
L	Length	meter	m
V	Potential	volt	V
p	Dipole moment	coulomb-meter	C·m
I	Current	ampere	A
J	Current density	ampere/meter^2	A/m^2
μ_e, μ_h	Mobility	meter^2/volt-second	m^2/V·s
e	Electronic charge	coulomb	C
σ	Conductivity	siemens/meter	S/m
R	Resistance	ohm	Ω
P	Polarization	coulomb/meter^2	C/m^2
$\chi_{e,m}$	Susceptibility		
C	Capacitance	farad	F
R_s	Sheet resistance	ohm per square	Ω
H	Magnetic field intensity	ampere/meter	A/m
K	Surface current density	ampere/meter	A/m
B	Magnetic flux density	tesla (or weber/meter^2)	T (or Wb/m^2)
μ_0, μ	Permeability	henry/meter	H/m
Φ	Magnetic flux	weber	Wb
V_m	Magnetic scalar potential	ampere	A
A	Vector magnetic potential	weber/meter	Wb/m
T	Torque	newton-meter	N·m
m	Magnetic moment	ampere-meter^2	A·m^2
M	Magnetization	ampere/meter	A/m
$\mathcal{R}$	Reluctance	ampere-turn/weber	A·t/Wb
L	Inductance	henry	H
M	Mutual inductance	henry	H
ω	Radian frequency	radian/second	rad/s
c	Velocity of light	meter/second	m/s
λ	Wavelength	meter	m
η	Intrinsic impedance	ohm	Ω
k	Wave number	meter^{-1}	m^{-1}
α	Attenuation constant	neper/meter	Np/m
β	Phase constant	radian/meter	rad/m
f	Frequency	hertz	Hz

TABLE B.1
(continued)

Symbol	Name	Unit	Abbreviation
$\mathcal{P}$	Poynting vector	watt/meter^2	W/m^2
P	Power	watt	W
δ	Skin depth	meter	m
Γ	Reflection coefficient		
s	Standing-wave ratio		
γ	Propagation constant	meter^{-1}	m^{-1}
G	Conductance	siemen	S
Z	Impedance	ohm	Ω
Y	Admittance	siemen	S
Q	Quality factor		

TABLE B.2
Conversion of International to gaussian and other units (use $c = 2.997\,924\,58 \times 10^8$)

Quantity	1 mks unit	= gaussian units	= other units
d	1 m	10^2 cm	39.37 in
F	1 N	10^5 dyne	0.2248 lb$_f$
W	1 J	10^7 erg	0.7376 ft-lb$_f$
Q	1 C	$10c$ statC	0.1 abC
ρ_v	1 C/m^3	$10^{-5}c$ statC/cm^3	10^{-7} abC/cm^3
D	1 C/m^2	$4\pi 10^{-3}c$ (esu)	$4\pi 10^{-5}$ (emu)
E	1 V/m	$10^4/c$ statV/cm	10^6 abV/cm
V	1 V	$10^6/c$ statV	10^8 abV
I	1 A	0.1 abA	$10c$ statA
H	1 A/m	$4\pi 10^{-3}$ oersted	$0.4\pi c$ (esu)
V_m	1 A·t	0.4π gilbert	$40\pi c$ (esu)
B	1 T	10^4 gauss	$100/c$ (esu)
Φ	1 Wb	10^8 maxwell	$10^6/c$ (esu)
A	1 Wb/m	10^6 maxwell/cm	
R	1 Ω	10^9 abΩ	$10^5/c^2$ statΩ
L	1 H	10^9 abH	$10^5/c^2$ statH
C	1 F	$10^{-5}c^2$ statF	10^{-9} abF
σ	1 S/m	10^{-11} abS/cm	$10^{-7}c^2$ statS/cm
μ	1 H/m	$10^7/4\pi$ (emu)	$10^3/4\pi c^2$ (esu)
ϵ	1 F/m	$4\pi 10^{-7}c^2$ (esu)	$4\pi 10^{-11}$ (emu)

TABLE B.3
Standard prefixes used with SI units

Prefix	Abbrev.	Meaning	Prefix	Abbrev.	Meaning
atto-	a-	10^{-18}	deka-	da-	10^{1}
femto-	f-	10^{-15}	hecto-	h-	10^{2}
pico-	p-	10^{-12}	kilo-	k-	10^{3}
nano-	n-	10^{-9}	mega-	M-	10^{6}
micro-	μ-	10^{-6}	giga-	G-	10^{9}
milli-	m-	10^{-3}	tera-	T-	10^{12}
centi-	c-	10^{-2}	peta-	P-	10^{15}
deci-	d-	10^{-1}	exa-	E-	10^{18}

APPENDIX C

MATERIAL CONSTANTS

Table C.1 lists typical values of the relative permittivity ϵ'_R or dielectric constant for common insulating and dielectric materials, along with representative values for the loss tangent. The values should only be considered representative for each material, and they apply to normal temperature and humidity conditions, and to very low audio frequencies. Most of them have been taken from "Reference Data for Radio Engineers,"[1] "The Standard Handbook for Electrical Engineers,"[2] and von Hippel,[3] and these volumes may be referred to for further information on these and other materials.

Table C.2 gives the conductivity for a number of metallic conductors, for a few insulating materials, and for several other materials of general interest. The values have been taken from the references listed previously, and they apply at zero frequency and at room temperature. The listing is in the order of decreasing conductivity.

Some representative values of the relative permeability for various diamagnetic, paramagnetic, ferrimagnetic, and ferromagnetic materials are listed in Table C.3. They have been extracted from the references listed above, and the

[1] See Suggested References for Chap. 11.

[2] See Suggested References for Chap. 5.

[3] von Hippel, A. R.: "Dielectric Materials and Applications," The Technology Press of the Massachusetts Institute of Technology, Cambridge, MA and John Wiley and Sons, Inc., New York, 1954.

TABLE C.1
ϵ'_R and ϵ''/ϵ'

Material	ϵ'_R	ϵ''/ϵ'
Air	1.0005	
Alcohol, ethyl	25	0.1
Aluminum oxide	8.8	0.000 6
Amber	2.7	0.002
Bakelite	4.74	0.022
Barium titanate	1200	0.013
Carbon dioxide	1.001	
Ferrite (NiZn)	12.4	0.000 25
Germanium	16	
Glass	4–7	0.002
Ice	4.2	0.05
Mica	5.4	0.000 6
Neoprene	6.6	0.011
Nylon	3.5	0.02
Paper	3	0.008
Plexiglas	3.45	0.03
Polyethylene	2.26	0.000 2
Polypropylene	2.25	0.000 3
Polystyrene	2.56	0.000 05
Porcelain (dry process)	6	0.014
Pyranol	4.4	0.000 5
Pyrex glass	4	0.000 6
Quartz (fused)	3.8	0.000 75
Rubber	2.5–3	0.002
Silica or SiO_2 (fused)	3.8	0.000 75
Silicon	11.8	
Snow	3.3	0.5
Sodium chloride	5.9	0.000 1
Soil (dry)	2.8	0.05
Steatite	5.8	0.003
Styrofoam	1.03	0.000 1
Teflon	2.1	0.000 3
Titanium dioxide	100	0.001 5
Water (distilled)	80	0.04
Water (sea)		4
Water (dehydrated)	1	0
Wood (dry)	1.5–4	0.01

data for the ferromagnetic materials is only valid for very low magnetic flux densities. Maximum permeabilities may be an order of magnitude higher.

Values are given in Table C.4 for the charge and rest mass of an electron, the permittivity and permeability of free space, and the velocity of light.[4]

[4] Cohen, E. R., and B. N. Taylor: "The 1986 Adjustment of the Fundamental Physical Constants," Pergamon Press, Elmsford, NY, 1986.

TABLE C.2
σ

Material	σ, S/m	Material	σ, S/m
Silver	6.17×10^7	Graphite	7×10^4
Copper	5.80×10^7	Silicon	2300
Gold	4.10×10^7	Ferrite (typical)	100
Aluminum	3.82×10^7	Water (sea)	5
Tungsten	1.82×10^7	Limestone	10^{-2}
Zinc	1.67×10^7	Clay	5×10^{-3}
Brass	1.5×10^7	Water (fresh)	10^{-3}
Nickel	1.45×10^7	Water (distilled)	10^{-4}
Iron	1.03×10^7	Soil (sandy)	10^{-5}
Phosphor bronze	1×10^7	Granite	10^{-6}
Solder	0.7×10^7	Marble	10^{-8}
Carbon steel	0.6×10^7	Bakelite	10^{-9}
German silver	0.3×10^7	Porcelain (dry process)	10^{-10}
Manganin	0.227×10^7	Diamond	2×10^{-13}
Constantan	0.226×10^7	Polystyrene	10^{-16}
Germanium	0.22×10^7	Quartz	10^{-17}
Stainless steel	0.11×10^7		
Nichrome	0.1×10^7		

TABLE C.3
μ_R

Material	μ_R
Bismuth	0.999 998 6
Paraffin	0.999 999 42
Wood	0.999 999 5
Silver	0.999 999 81
Aluminum	1.000 000 65
Beryllium	1.000 000 79
Nickel chloride	1.000 04
Manganese sulfate	1.000 1
Nickel	50
Cast iron	60
Cobalt	60
Powdered iron	100
Machine steel	300
Ferrite (typical)	1000
Permalloy 45	2500
Transformer iron	3000
Silicon iron	3500
Iron (pure)	4000
Mumetal	20 000
Sendust	30 000
Supermalloy	100 000

TABLE C.4
Physical Constants

Quantity	Value
Electron charge	$e = (1.602\,177\,33 \pm 0.000\,000\,46) \times 10^{-19}$ C
Electron mass	$m = (9.109\,389\,7 \pm 0.000\,005\,4) \times 10^{-31}$ kg
Permittivity of free space	$\epsilon_0 = 8.854\,187\,817 \times 10^{-12}$ F/m
Permeability of free space	$\mu_0 = 4\pi 10^{-7}$ H/m
Velocity of light	$c = 2.997\,924\,58 \times 10^{8}$ m/s

APPENDIX D

ORIGINS OF THE COMPLEX PERMITTIVITY

As we learned in Chap. 5, a dielectric can be modeled as an arrangement of atoms and molecules in free space, which can be polarized by an electric field. The field forces positive and negative bound charges to separate against their Coulomb attractive forces, thus producing an array of microscopic dipoles. The molecules can be arranged in an ordered and predictable manner (such as in a crystal) or may exhibit random positioning and orientation, as would occur in an amorphous material or a liquid. The molecules may or may not exhibit permanent dipole moments (existing before the field is applied), and if they do, they will usually have random orientations throughout the material volume. As discussed in Sec. 5.7, the displacement of charges in a regular manner, as induced by an electric field, gives rise to a macroscopic polarization, **P**, defined as the dipole moment per unit volume:

$$\mathbf{P} = \lim_{\Delta v \to 0} \frac{1}{\Delta v} \sum_{i=1}^{N \Delta v} \mathbf{p_i} \tag{D.1}$$

where N is the number of dipoles per unit volume and $\mathbf{p_i}$ is the dipole moment of the ith atom or molecule, found through

$$\mathbf{p_i} = Q_i \mathbf{d}_i \tag{D.2}$$

Q_i is the positive one of the two bound charges composing dipole i, and $\mathbf{d}_i$ is the distance between charges, expressed as a vector from the negative to the positive

charge. Again, borrowing from Sec. 5.7, the electric field and the polarization are related through

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E} \tag{D.3}$$

where the electric susceptibility, χ_e, forms the more interesting part of the dielectric constant:

$$\epsilon_R = 1 + \chi_e \tag{D.4}$$

Therefore, to understand the nature of ϵ_R, we need to understand χ_e, which in turn means that we need to explore the behavior of the polarization, **P**.

Here, we consider the added complications of how the dipoles respond to a time-harmonic field that propagates as a wave through the material. The result of applying such a forcing function is that *oscillating* dipole moments are set up, and *these in turn establish a polarization wave that propagates through the material*. The effect is to produce a polarization function, $\mathbf{P}(z, t)$, having the same functional form as the driving field, $\mathbf{E}(z, t)$. The molecules themselves do not move through the material, but their oscillating dipole moments collectively exhibit wave motion, just as waves in pools of water are formed by the up and down motion of the water. From here, the description of the process gets complicated and in many ways beyond the scope of our present discussion. We can form a basic qualitative understanding, however, by considering the classical description of the process, which is that the dipoles, once oscillating, behave as microscopic antennas, re-radiating fields that in turn co-propagate with the applied field. Depending on the frequency, there will be some phase difference between the incident field and the radiated field at a given dipole location. This results in a net field (formed through the superposition of the two) that now interacts with the next dipole. Radiation from this dipole adds to the previous field as before, and the process repeats from dipole to dipole. The accumulated phase shifts at each location are manifested as a net slowing down of the phase velocity of the resultant wave. Attenuation of the field may also occur which, in this classical model, can be accounted for by partial phase cancellation between incident and radiated fields.

In our classical model, the medium is an ensemble of identical fixed electron oscillators, in which the Coulomb binding forces on the electrons are modeled by springs that attach the electrons to the positive nuclei. We consider electrons for simplicity, but similar models can be used for any bound charged particle. Figure D.1 shows a single oscillator, located at position z in the material, and oriented along x. A uniform plane wave, assumed linearly polarized along x, propagates through the material in the z direction. The electric field in the wave displaces the electron of the oscillator in the x direction through a distance represented by the vector **d**; a dipole moment is thus established,

$$\mathbf{p}(z, t) = -e\mathbf{d}(z, t) \tag{D.5}$$

where the electron charge, e, is treated as a positive quantity. The applied force is

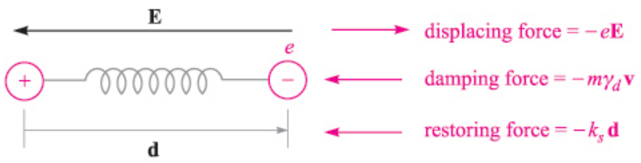


FIGURE D.1
Atomic dipole model, with Coulomb force between positive and negative charge modeled by that of a spring having spring constant, k_s. An applied electric field displaces the electron through distance d, resulting in dipole moment, $\mathbf{p} = -e\mathbf{d}$.

$$\mathbf{F}_a(z, t) = -e\mathbf{E}(z, t) \tag{D.6}$$

We need to remember that $\mathbf{E}(z, t)$ at a given oscillator location is the *net* field, composed of the original applied field plus the radiated fields from all other oscillators. The relative phasing between oscillators is precisely determined by the spatial and temporal behavior of $\mathbf{E}(z, t)$.

The restoring force on the electron, $\mathbf{F}_r$, is that produced by the spring which is assumed to obey Hooke's law:

$$\mathbf{F}_r(z, t) = -k_s\mathbf{d}(z, t) \tag{D.7}$$

where k_s is the spring constant (not to be confused with the propagation constant). If the field is turned off, the electron is released and will oscillate about the nucleus at the *resonant frequency*, given by

$$\omega_0 = \sqrt{k_s/m} \tag{D.8}$$

where m is the mass of the electron. The oscillation, however, will be damped since the electron will experience forces and collisions from neighboring oscillators. We model these as a velocity-dependent damping force:

$$\mathbf{F}_d(z, t) = -m\gamma_d\mathbf{v}(z, t) \tag{D.9}$$

where $\mathbf{v}(z, t)$ is the electron velocity. Associated with this damping is the *dephasing* process among the electron oscillators in the system. Their relative phasing, once fixed by the applied sinusoidal field, is destroyed through collisions and dies away exponentially until a state of totally random phase exists between oscillators. The $1/e$ point in this process occurs at the *dephasing time* of the system, which is inversely proportional to the damping coefficient, γ_d (in fact it is $2/\gamma_d$). We are, of course, driving this damped resonant system with an electric field at frequency ω. We can therefore expect the response of the oscillators, measured through the magnitude of $\mathbf{d}$, to be frequency-dependent in much the same way as an RLC circuit is when driven by a sinusoidal voltage.

We can now use Newton's second law, and write down the forces acting on the single oscillator of Fig. D.1. To simplify the process a little we can use the complex form of the electric field:

$$\mathbf{E}_c = \mathbf{E}_0 e^{-jkz} e^{j\omega t} \tag{D.10}$$

Defining **a** as the acceleration vector of the electron, we have

$$m\mathbf{a} = \mathbf{F}_a + \mathbf{F}_r + \mathbf{F}_d$$

or

$$m\frac{\partial^2 \mathbf{d}_c}{\partial t^2} + m\gamma_d \frac{\partial \mathbf{d}_c}{\partial t} + k_s \mathbf{d}_c = -e\mathbf{E}_c \tag{D.11}$$

Note that since we are driving the system with the complex field, $\mathbf{E}_c$, we anticipate a displacement wave, $\mathbf{d}_c$, of the form:

$$\mathbf{d}_c = \mathbf{d}_0 e^{-jkz} e^{-j\omega t} \tag{D.12}$$

With the waves in this form, time differentiation produces a factor of $j\omega$. Consequently (D.11) can be simplified and rewritten in phasor form:

$$-\omega^2 \mathbf{d}_s + j\omega\gamma_d \mathbf{d}_s + \omega_0^2 \mathbf{d}_s = -\frac{e}{m}\mathbf{E}_s \tag{D.13}$$

where (D.4) has been used. We now solve (D.13) for $\mathbf{d}_s$, obtaining

$$\mathbf{d}_s = \frac{-(e/m)\mathbf{E}_s}{(\omega_0^2 - \omega^2) + j\omega\gamma_d} \tag{D.14}$$

The dipole moment associated with displacement $\mathbf{d}_s$ is

$$\mathbf{p}_s = -e\mathbf{d}_s \tag{D.15}$$

The polarization of the medium is then found assuming that all dipoles are identical. Eq. (D.1) thus becomes

$$\mathbf{P}_s = N\mathbf{p}_s$$

which, when using (D.14) and (D.15), becomes

$$\mathbf{P}_s = \frac{Ne^2/m}{(\omega_0^2 - \omega^2) + j\omega\gamma_d}\mathbf{E}_s \tag{D.16}$$

Now, using (D.3) we identify the susceptibility associated with the resonance as

$$\chi_{res} = \frac{Ne^2}{\epsilon_0 m}\frac{1}{(\omega_0^2 - \omega^2) + j\omega\gamma_d} = \chi'_{res} - j\chi''_{res} \tag{D.17}$$

The real and imaginary parts of the permittivity are now found through the real and imaginary parts of χ_{res}: Knowing that

$$\epsilon = \epsilon_0(1 + \chi_{res}) = \epsilon' - j\epsilon''$$

we find

$$\epsilon' = \epsilon_0(1 + \chi'_{res}) \tag{D.18}$$

and

$$\epsilon'' = \epsilon_0 \chi''_{res} \tag{D.19}$$

The above expressions can now be used in Eqs. (35) and (36) in Chap. 11 to evaluate the attenuation coefficient, α, and phase constant, β, for the plane wave as it propagates through our resonant medium.

The real and imaginary parts of χ_{res} as functions of frequency are shown in Fig. D.2 for the special case in which $\omega \doteq \omega_0$. Eq. (D.17) in this instance becomes

$$\chi_{res} \doteq -\frac{Ne^2}{\epsilon_0 m \omega_0 \gamma_d}\left(\frac{j+\delta_n}{1+\delta_n^2}\right) \tag{D.20}$$

where the *normalized detuning* parameter, δ_n, is

$$\frac{2}{\gamma_d}(\omega - \omega_0) \tag{D.21}$$

Key features to note in Fig. D.2 include the symmetric χ_e'' function, whose full-width at its half-maximum amplitude is γ_d. Near the resonant frequency, where χ_{res}'' maximizes, wave attenuation maximizes as seen from Eq. (35), Chap. 11. Additionally, we see that away from resonance, attenuation is relatively weak, and the material becomes transparent. As Fig. D.2 shows, there is still significant variation of χ_{res}' with frequency away from resonance, which leads to a frequency-dependent refractive index; this is expressed approximately as

$$n \doteq \sqrt{1+\chi_{res}'} \qquad \text{(away from resonance)} \tag{D.22}$$

This frequency-dependent n, arising from the material resonance, leads to phase and group velocities that also depend on frequency. Thus, group dispersion, leading to pulse broadening effects as discussed in Chap. 12, can be directly attributable to material resonances.

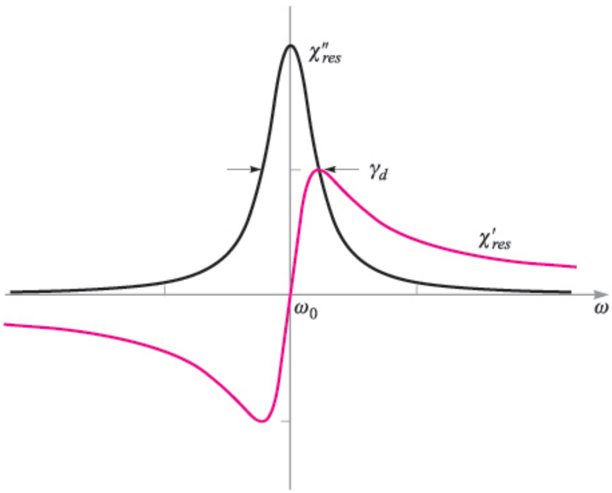


FIGURE D.2
Plots of the real and imaginary parts of the resonant susceptibility, χ_{res}, as given by Eq. (D.20). The full-width at half-maximum of the imaginary part, χ_{res}'', is equal to the damping coefficient, γ_d.

Somewhat surprisingly, the classical "spring model" described here can provide very accurate predictions on dielectric constant behavior with frequency (particularly off-resonance) and can be used to a certain extent to model absorption properties. The model is insufficient, however, when attempting to describe the more salient features of materials; specifically, it assumes that the oscillating electron can assume any one of a continuum of energy states, when in fact energy states in any atomic system are quantized. As a result, the important effects arising from transitions between discrete energy levels, such as spontaneous and stimulated absorption and emission, are not included in our classical spring system. Quantum mechanical models must be used to fully describe the medium polarization properties, but the results of such studies often reduce to those of the spring model when field amplitudes are very low.

Another way that a dielectric can respond to an electric field is through the orientation of molecules that possess permanent dipole moments. In such cases, the molecules must be free to move or rotate, and so the material is typically a liquid or a gas. Figure D.3 shows an arrangement of polar molecules in a liquid (such as water) in which there is no applied field (D.3*a*) and where an electric field is present (D.3*b*). Applying the field causes the dipole moments, previously having random orientations, to line up, and so a net material polarization, **P**, results. Associated with this, of course, is a susceptibility function, χ_e, through which **P** relates to **E**.

Some interesting developments occur when the applied field is time-harmonic. With field periodically reversing direction, the dipoles are forced to follow, but do so against their natural propensity to randomize, owing to thermal motion. Thermal motion thus acts as a "restoring" force, effectively opposing the applied field. We can also think of the thermal effects as viscous forces that introduce some difficulty in "pushing" the dipoles back and forth. One might expect (correctly) that polarizations of greater amplitude in each direction can be attained at lower frequencies, since enough time is given during each cycle for the dipoles to achieve complete alignment. The polarization amplitude will weaken as the frequency increases, since there is no longer enough time for complete alignment during each cycle. This is the basic description of *dipole relaxation* mechanism for the complex permittivity. There is no resonant frequency associated with the process.

The complex susceptibility associated with dipole relaxation is essentially that of an "overdamped" oscillator, and is given by

$$\chi_{rel} = \frac{Np^2/\epsilon_0}{3k_B T(1 + j\omega\tau)} \tag{D.23}$$

where p is the permanent dipole moment magnitude of each molecule, k_B is Boltzmann's constant, and T is the temperature in degees Kelvin. τ is the thermal randomization time, defined as the time for the polarization, **P**, to relax to $1/e$ of its original value when the field is turned off. χ_{rel} is complex, and so it will possess absorptive and dispersive components (imaginary and real parts) as we

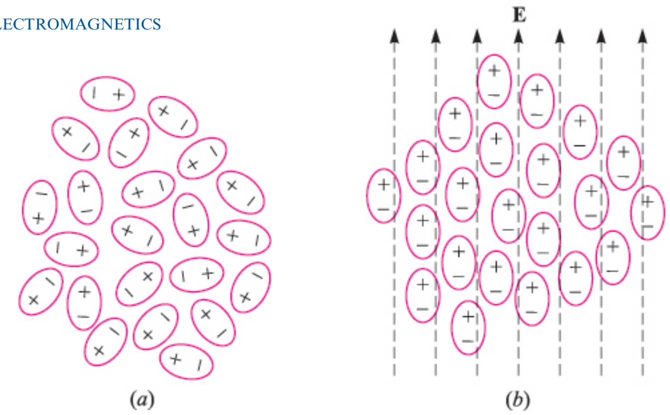


FIGURE D.3
Idealized sketches of ensembles of polar molecules under conditions of (*a*) random orientation of the dipole moments, and (*b*) dipole moments aligned under the influence of an applied electric field. Conditions in (*b*) are greatly exaggerated, since typically only a very small percentage of the dipoles align themselves with the field. But still enough alignment occurs to produce measurable changes in the material properties.

found in the resonant case. The form of Eq. (D.23) is identical to that of the response of a series RC circuit driven by a sinusoidal voltage (where τ becomes RC).

Microwave absorption in water occurs through the relaxation mechanism in polar water molecules, and is the primary means by which microwave cooking is done, as discussed in Chap. 11. Frequencies near 2.5 GHz are typically used, since these provide the optimum penetration depth. The peak water absorption arising from dipole relaxation occurs at much higher frequencies, however.

A given material may possess more than one resonance and may have a dipole relaxation response as well. In such cases, the net susceptibility is found in frequency domain by the direct sum of all component susceptibilities. In general, we may write:

$$\chi_e = \chi_{rel} + \sum_{i=1}^{n} \chi_{res}^{i} \qquad \text{(D.24)}$$

where χ_{res}^{i} is the susceptibility associated with the ith resonant frequency, and n is the number of resonances in the material. The reader is referred to the suggested references for Chap. 11 for further reading on resonance and relaxation effects in dielectrics.

INDEX

A (*see* Potential, vector magnetic)
Absolute potential, 92
Admittance, 459, 462
Ampere, defined, 120, 535
Ampere's circuital law, 226, 232–239, 249, 312, 331, 332
  point form of, 246, 263, 330, 334, 337, 516
Ampere-turn, 300
Angular dispersion, 422
Anisotropic material, 126, 142–143, 291, 296
Antenna patterns, 519–520
Antennas, 514–524
  dipole, 522
  monopole, 523
Antiferromagnetic material, 291–292
Area
  of parallelogram, 14
  of vector surface, 10
Attenuation coefficient, 357, 358

B, defined, 251
Biot-Savart law, 225–232. 236, 237
Bound charge, 137–143, 146, 292, 544–546
Bound (amperian) current, 293–296, 331
Boundary conditions
  conductor-dielectric, 148–149, 169, 336–337
  conductor-free space, 134, 336–337
  dielectric-dielectric, 144–149, 336–337
  magnetic, 297–299, 336–337
Brewster angle, 419–420

Capacitance
  between coaxial cylinders, 154, 173, 203
  between concentric spheres, 205
  between cone and plane, 207
  by current analogy, 183
  from curvilinear squares, 173
  between cylinder and plane, 159, 160
  defined, 151, 308
  between parallel cylinders, 162
  between parallel planes, 151, 152, 155–157
  partial, 153
  of *pn* junction, 210
Capacitor
  coaxial, 112, 154, 173, 203
  energy stored in, 153
  multiple dielectric, 155–157

Capacitor (*Continued*)
parallel-plate, 151–152, 155–157, 202
spherical, 154–155, 205
Cartesian components, 6, 19, 22
Cartesian coordinate system, 4–15, 72, 88, 104, 197, 242, 530
for expressing curl, 242
for expressing divergence, 72
for expressing gradient, 104
Laplacian, 197
transformation to other coordinate systems, 18–19, 22–23
Cgs system of units, 536
Characteristic impedance, 440
Charge, 28, 53, 120, 150
bound, 137–143, 146, 292
conservation of, 122
on electron, 543
forces on (*see* Force, on charge)
free, 141, 146, 294–296
magnetic, 252, 307
point, 28–31, 59, 92, 106, 197
E field of, 31–33
potential field of, 92–97
Charge density
line, 38, 59, 97, 112, 197
E field of, 40, 42
potential field of, 92, 97
surface, 44, 59, 97, 129, 144, 150, 197
E field of, 45, 202
potential field of, 97, 202
volume, 35, 59, 70, 97, 120, 141, 197
Chirped pulse, 429
Circuit(s)
electric, 98, 332
magnetic, 299–308
substitution of, 328
Circulation, 243, 329
Closed path, 98, 233
Closed surface, 57–59, 249–250
Coaxial transmission line
E field of, 203
H field of, 235–236
R, *G*, *L*, and *C* of, 442–445
Coercive force, 303
Components
cartesian, 6, 19, 22
normal (*see* Normal component)
transformations of, 18–19, 22–23
vector, 3–4, 6, 12
Conduction, 120, 124
in metals, 120, 124–128, 150, 276
skin effect, 369–372
Conduction current, 126
Conductivity, 126
defined, 126
magnetic analog for, 301
of semiconductors, 136
table of values, 541, 542
Conductor(s)
filamentary, 120, 225, 312
metallic (*see* Conduction, in metals)
moving, 323–329
parallel, force between, 281–282
perfect, 132
Conservation of charge, 122
Conservative field, 95–99, 257
Continuity equation, 122, 123, 148, 226, 330
Coordiate axes, 5, 6, 16, 20
Coordinate system
cartesian (*see* Cartesian coordinate system)
cylindrical (*see* Cylindrical coordinate system)
general curvilinear, 529–532
rectangular (*see* Cartesian coordinate system)
right-handed, 5, 14, 17, 20
spherical (*see* Spherical coordinate system)
transformation between, 18–19, 22–23
Coplanar vectors, 3
Coulomb, 29, 120
Coulomb's law, 28, 83, 226, 232, 276, 288
Critical angle, 418
Cross product, 13–15
Cutoff frequency, 493, 503, 504
Cutoff wavelength, 493, 513
Curie temperature, 291, 292
Curl, 239–246, 329, 340
defined, 242
Current (*I*), 120, 225, 300
amperian (bound), 293–296, 331
conduction, 126
continuity of, 122
convection, 121
defined, 121
displacement, 329–333
filamentary, 120
line, **H** field of, 229, 231

Current density, 120, 227, 300, 311
  conduction, 126, 182, 331, 334
  convection, 121, 331, 334
  defined, 121
  displacement, 330–333
  surface, 227, 337
Current element, differential, 225, 514
  force on, 276–279
Current loop, differential, 284–286
  torque on, 285
Current sheet, **H** field of, 237
Curvilinear coordinates, 23, 529–532
Curvilinear squares, 170–175, 183, 229, 257
Cylinder(s)
  capactance of, to plane, 159
  circular, 15
  parallel, capacitance between, 162
Cylindrical coordinate system, 15–19, 72, 88, 104, 197, 243, 530
  for expressing curl, 243
  for expressing divergence, 72
  for expressing gradient, 88
  Laplacian, 197
  transformation between, 18–19

D, defined, 56
∇, del operator, 74
Determinant, 14, 242
Diamagnetic material, 288, 296
Dielectric, 54, 137–149, 182
  perfect, 144–148
Dielectric constant, 142, 545 (*See also* Permittivity)
  table of values, 541
Dielectric hysteresis, 142
Dielectric interface, 144–148, 388–390
Dielectric waveguide
  cutoff condition for, 512–513
  guided modes in, 509–511
  leaky waves in, 507–508
  plane wave analysis, 507–510
  surface waves in, 510
  transverse resonance in, 512
Differential distance, 6, 17, 20, 89
Differential surface, 6, 17, 21, 59
Differential volume, 6, 17, 21
Dipole
  E field of, 108
  electric, 106, 139, 285, 544–546
  magnetic, 285
  point, 108
  potential field of, 107, 109
  spring model of, 546
Dipole antenna, 522
Dipole moment, 109, 139, 544
  magnetic, 285–287
  per unit volume, 139, 544
Directed distance, 8, 30
Direction of vector, 6, 42, 46
Dispersion
  angular, 422
  group velocity, 421–430
  parameter, 428
Displacement current, 329–333
Displacement flux, 54
Displacement flux density, 55
Distributed circuit, 436
Divergence, 70–72, 123, 169, 340
  defined, 71
Divergence theorem, 76–78
Domain, ferromagnetic, 290
Dot operation, 74
Dot product, 10–12, 19, 22
Drift velocity, 125

E, defined, 32
Effective impedance, 415–416
Electric circuit, 98, 332
Electric dipole, 106, 139, 285
Electric field, energy density in, 110–114
Electric field intensity, 31–35, 84, 99, 141–143, 275
  defined, 32
  of dipole, 108
  of line charge, 40, 42, 89
  magnitude of, 101
  motional, 327
  of n point charges, 34
  of point charge, 31
  of radial planes, 204
  of sheet of charge, 45
  of two sheets of charge, 45
Electric flux, 53–54, 58
Electric flux density, 54, 130, 141–144, 170
Electric susceptibility, 142, 545–546
Electrolytic trough, 183

Electromotive force (emf), 300, 323–329
motional, 327
Electron, 29, 124, 138, 207, 543
charge on, 543
conduction, 125
free, 125
mass of, 543
radius of, 29
Electron mobility, 125
Electron orbit, 124, 288–291
Electron spin, 288–291
Electrostatic fields, 98, 129, 134, 169, 300
Electrostatic potential, 91, 92, 254, 337
Element, volume, 6, 17, 21
Emu system of units, 536, 538
Energy (*See also* Work)
in gravitational field, 186
kinetic, 124
to move point charge, 84
potential (*see* Potential energy)
quantum of, 124
stored in capacitor, 153
stored in inductor, 310
thermal, 126, 136
Energy density, 114, 307
in electric field, 110–114, 366
in magnetic field, 306–307, 366
English system of units, 29, 538–539
Equipotential surface, 95, 101, 134, 150, 170
Esu system of units, 536, 538–539
External inductance, 313

Fabry-Perot interferometer, 404–405
Farad, 29, 151
defined, 151
Faraday's law, 307, 323–329, 334, 335, 336, 338
Ferrimagnetic material, 288, 291, 292, 358
Ferrite, 292, 295, 358, 541
Ferroelectric material, 142
Ferromagnetic domain, 290–291
Ferromagnetic material, 289–291, 296, 302–305
Field(s), 2
conservative, 95–99
constant, 45
electrostatic, 98, 129, 134, 169, 300
flux-density type of, 54–55, 71
force-type, 31, 55
gravitational, 84, 93, 95, 99, 185–187
inverse-cube, 109
inverse distance, 42, 95, 97
inverse-square law (*see* Inverse square law field)
magnetic (*see* Magnetic field)
non-conservative, 99 potential (*see* Potential field)
scalar, 2, 95, 103
sketches of, 46–49
vector, 2–4, 9
Field map, 46, 170
Filamentary conductor, 120, 225, 312
Filamentary current, 120, 231
Fluid dynamics, 187, 243
Fluid-flow maps, 187
Flux
electric, 53–54, 58, 171
fringing, 305
leakage, 305
magnetic (*see* Magnetic flux)
Flux density
displacement, 55
electric, 54, 130, 141–144, 170
magnetic (*see* Magnetic flux density)
remnant, 303–304
Flux-density type of field, 54–55, 71
Flux line, 53–55
Flux linkage, 308, 329
Force, 10
on charge, 28–31, 84, 125, 275, 331, 335
on closed circuit, 282
coercive, 303
on conductor, 278
on differential current element, 276–282
between differential current elements, 280–282
Lorentz, 275, 335
on magnetic material, 306–308
moment of a, 283
on moving charge, 275
between parallel conductors, 281–282
Force-type field, 31, 55
Fourier transform, 426
Fourier series, 217
Free charge, 141, 146, 294–296

Free electron, 125
Free space
  permeability of, 251, 535, 543
  permittivity of, 29, 536, 543
Free spectral range, 409
Fringing flux, 305

Gauss, defined, 251
Gauss' law, 57–59, 67, 83, 130
  applications of, 62–70
  for the magnetic field, 252, 336
  point form of, 73, 143, 196
Gaussian surface, 59
Gaussian system of units, 536, 538–539
Gradient, 102, 169, 531
  defined, 102
  potential, 99, 254
Gravitational field, 84, 93, 95, 99, 185–187
Gravity, acceleration due to, 186
Ground, 93
Group delay, 426, 427
Group velocity, 425
Group velocity dispersion, 425

H, defined, 225, 294
Half-wave matching, 403–404
Hall effect, 276
Heaviside-Lorentz system of units, 536
Helmholtz equation, 352
Henry, defined, 251, 309
Homogeneous material, 150, 172, 255
Hydraulics, 188, 243
Hysteresis
  dielectric, 142
  magnetic, 290, 303

I (*see* Current)
Identities, vector, 532–533
Images, 106, 134–136
Impedance
  characteristic, 440
  effective, 415–416
  input, 403, 441
  intrinsic, 355, 359, 360, 364, 373
  normalized, 453
  transformation of, 406–408

Incidence
  normal, 388–394, 400–408
  oblique, 411–421
Inductance, 308–314
  defined, 309
  external, 313
  internal, 313
  mutual, 313–314
  of solenoid, 314
  of toroid, 309
Input impedance, 403
Insulator, 125, 137, 541 (*See also* Dielectric)
Integral
  line (*see* Line integral)
  surface, 59, 76
    closed, 59, 76
  volume, 36, 76
Internal inductance, 313
International System of Units, 28, 534–539
Intrinsic impedance, 355, 359, 360, 364, 373
Inverse square law field, 32, 42, 62, 95, 109
Isotropic material, 126, 142, 291, 296
Iteration method, 176–182

J, defined, 120
Joule, 31

K, defined, 227
Kelvin, defined, 535
Kilogram, defined, 535
Kirchhoff's voltage law, 99, 145, 301, 438

L, defined, 309
Laplace's equation, 176, 195–207, 211–219, 255
Laplacian, 197, 532
  defined, 197
  of a vector, 264–266
Leakage flux, 305
Lenz's law, 323
Light, velocity of, 341, 353, 535, 543
Line charge
  E field of, 40. 42, 89
  parallel, potential field of, 157–162
  potential field of, 92, 97

Line current, **H** field of, 229, 231
Line integral, 85
  closed, 98, 241, 247, 324, 327
  path of, 85–93
Linearity, 31, 126, 142, 297
Linkage, flux, 308, 329
Lorentz force, 335
Loss tangent, 358, 362–365, 369, 373, 541
Lumped elements, 435–436

M, defined, 293
M, defined, 313
m, defined, 285
Magnet, permanent, 225
Magnetic boundary conditions, 297–299, 336–337
Magnetic charge, 252, 307, 334
Magnetic circuit, 299–308
Magnetic dipole, 285
Magnetic dipole moment, 285–287
Magnetic field
  energy density in, 307
  Gauss' law for, 252, 336
Magnetic field intensity, 225, 294
  of coaxial cable, 234–236
  of current sheet, 237
  defined, 252, 336
  of finite line current, 231
  of line current, 229, 231
  of solenoid, 238–239
  of toroid, 239
Magnetic flux, 251–254, 301, 312, 323
  defined, 252
Magnetic flux density, 10, 14, 252–254, 300
  defined, 251
  remnant, 303, 304
Magnetic imaging, nuclear, 288
Magnetic interface, 297–299
Magnetic materials, 287–292
  characteristics of, table, 292
Magnetic moment, 285–292
Magnetic pole, 334
Magnetic potential
  scalar, 254–257, 300–301
    defined, 254
  vector, 257–261, 311, 337–341
    defined, 257–258
Magnetic susceptibility, 295
Magnetization, 292–299
Magnetization curve, 303–304
Magnetohydrodynamics, 275
Magnetomotive force (mmf), 300–305
Magnetostriction, 291
Magnitude, vector (*see* Vector magnitude)
Map
  curvilinear square, 170–175, 183, 229, 257
  field, 46, 170
  fluid-flow, 187
Mass, 186
Material(s) (*See also* Conduction in metals; Dielectric; Semiconductor)
  anisotropic, 126, 142, 291, 196
  antiferromagnetic, 291–292
  diamagnetic, 288, 296
  ferrimagnetic, 288, 291, 358
  ferroelectric, 142
  ferromagnetic, 289–292, 296, 302–305
  homogeneous, 150, 172, 255
  isotropic, 126, 142, 291, 296
  magnetic, 287–292
  nonmagnetic, 296
  paramagnetic, 290, 292, 296
  superparamagnetic, 292
Maxwell's equations
  in differential form, 333–334
  in integral form, 335–336
  non-time-varying, 73, 246, 252
  in phasor form, 351
  in sourceless media, 349
  time-varying, 325–337
Meter, defined, 534–535
Mho, defined, 126
Microwave oven, 361
MKS units, 534–539
Mmf (magnetomotive force), 300–305
Mobility, 125
Molecule
  nonpolar, 139
  polar, 139, 544
Moment
  dipole, 109, 139, 544
    per unit volume, 139, 544
  of a force, 283
    magnetic dipole, 285–287
Monopole antenna, 523
Multiplication
  vector

cross, 13–15
dot, 10–12, 18–19, 22
vector by scalar, 4
vector by vector, 4, 10, 13
Multipole, 10

Neper, 357
Newton, defined, 29, 535
Noise, 36
Nonconservative field, 99
Nonlinearity, 142, 143, 300
Nonmagnetic material, 296
Nonpolar molecule, 139
Normal component
at conductor boundary, 131
at dielectric boundary, 145
at magnetic boundary, 298
at perfect conductor, 336
Normal incidence, 388
Normal to surface, 11, 13, 45, 58, 102
Normalized admittance, 462
Normalized impedance, 453
Nuclear spin, 288, 292
Nuclear magnetic imaging, 288

Ohm
defined, 128
per square, 185
Ohm's law, 128
point form of, 126, 137, 182, 300, 336
Omega-beta (ω-β) diagram, 423, 427
Operator
del, 74
scalar, 75
vector, 74
Optical fiber waveguide:
structure of, 486, 513
cutoff condition for, 513

P, defined, 139, 544
p, defined, 109, 544, 545
$\mathcal{P}$, defined, 367
Paddle wheel, 244
Paper, conducting, 184
Parallel conductors, force between, 281–282
Parallel cylinders, capacitance between, 162
Parallel line charges, potential field of, 157–162
Parallel-plate capacitor, 151–152, 155–157
Parallel-plate transmission line, 446–448
Parallel-plate waveguide
cutoff conditions for, 493
as distinguished from transmission line, 487
group velocity in, 496
mode field expressions, 499–500
phase constant of, 489
phase velocity in, 496
plane wave analysis of, 488–497
transverse resonance in, 490–491
wave equation applied to, 497–501
Parallelepiped, rectangular, 6
Parallelogram, area of, 14
Parallelogram law, 3
Paramagnetic material, 290, 292, 296
Partial capacitance, 153
Penetration, depth of, 361, 371
Perfect conductor, 132, 183, 336
Perfect dielectric, 144–148
Permanent magnet, 225
Permeability, 292–296
of free space, defined, 251, 535, 543
relative, 295
table of values, 542
Permittivity
complex, 544–551
defined, 142–143
of free space, 29
defined, 536, 543
relative, 142
table of values, 541
Phase constant, 353
Phase velocity, 353–354, 359, 410
Phasor, 2, 349–351
Φ, defined, 252
Plane of incidence, 412
Plane waves, uniform
attenuation of (*see* Plane waves in lossy media)
defined, 355
in dielectrics (*see* Plane waves in lossy media)
dispersion of (*see* Dispersion)
group velocity of, 425

Plane waves, uniform (*Continued*)
phase constant of, 353
phase velocity of, 353–354, 359
phasor form of, 350, 352
polarization of (*see* Polarization, plane wave)
power in, 365–369
Poynting vector, 367, 369
propagation in general directions, 408–411
real instantaneous form of, 353
reflection of (*see* Reflection, plane wave)
in waveguides, 488–497, 504, 507–510
wavelength of, 353, 359
Plane waves in lossy media
approximations
small loss tangent, 362–365
large loss tangent, 369–373
attenuation coefficient for, 357, 358
complex permittivity, 358
complex propagation constant for, 357
loss tangent, 358, 362–365, 369, 373
Point charge, 28–31, 59, 92, 106, 197
E field of, 31, 32
potential field of, 93–95
Point charges
E field of, 34
potential field of, 96
Poisson's equation, 196, 207–211
Polarization, dielectric, 139–144, 147, 335, 544–547
Polarization, plane wave, 376–382,
circular, 379–382
defined, 377
elliptical, 378–379
linear, 377
p-polarization, 412
s-polarization, 413
Pole, magnetic, 334
Potential
absolute, 92
electrostatic, 91, 92, 254, 337
retarded, 337–342, 515–516
scalar magnetic, 254–257. 300–301
vector magnetic, 257–261, 311, 337–341
zero reference for, 93
Potential difference, 91–94, 128, 150, 300
defined, 91
Potential energy, 99, 110–114, 124, 139, 306
Potential field, 99
of concentric spheres, 205
of cones, 206
of dipole, 107, 109
of line charge, 92, 97
of *n* point charges, 96
of parallel line charges, 157–162
of point charge, 93–95
of radial planes, 204
Potential gradient, 99, 254
Power, 365–369
Power series, 213
Poynting theorem, 366
Poynting vector
instantaneous, 367
time-average, 369
Prism, as dispersive element, 422
Product solution, 211
Projection
scalar, 12
vector, 12, 22
Propagation constant
for transmission line, 439
for uniform plane wave, 357
Pulse, electric field of, 426
Pulse spectrum, 426–427
Pulse spreading, 426–429

Quantum of energy, 124
Quantum theory, 124, 288
Quarter-wave matching, 405–406

Radiation resistance, 520–521
Radome, 404
Rationalized mks units, 536
Rectangular coordinate system (*see* Cartesian coordinate system)
Rectangular waveguide
cutoff conditions for, 503, 504
mode field expressions for, 502, 504
plane waves in, 504
Reference, potential, 93
Reflected power, 394
Reflection, plane waves
at multiple interfaces, 400–408
at a single interface
at normal incidence, 388–390
at oblique incidence, 411–416

Reflection, transmission line waves, 440–441
Reflection coefficient
  for plane waves at normal incidence, 390, 403
  for plane waves at oblique incidence, 415
  for transmission line waves, 440–441
Refractive index, defined, 412
Relative permeability, 295
  table of values, 542
Relative permittivity, 142
  table of values, 541
Relativity, 329
Relaxation method, 182
Relaxation time, 149
Reluctance, 301
Remnant flux density, 303–304
Resistance, 128, 183–185, 301
  defined, 128
  radiation, 520–521
  sheet, 184
Resistivity, defined, 126
Retarded potentials, 337–342
Right-hand rule, 14, 20
Right-handed coordinate system, 5, 14, 17, 19
Right-handed screw, 5, 13, 14
Rounding off, 178

Scalar field, 2, 95, 103
Scalar magnetic potential, 254–257, 300–301
  defined, 254
Scalar operator, 75
Scalar product, 10–12, 19, 22
Scalar projection, 12
Semiconductor, 125, 136–137, 207
  intrinsic, 136
  n-type, 138, 207
  p-type, 138, 207
  pn junction, 207
Separation constant, 212
Series, 34, 213, 217
Sheet of charge, field of, 44
Sheet resistance, 184
Shielding, 66, 236
Siemen, defined, 126
Sink, 71
Skin depth, 371
Skin effect, 369–376
Smith chart
  derivation of, 452–458
  picture of, 458
  as admittance chart, 461–463
  standing wave ratio from, 459
  single-stub matching with, 461–463
Snell's law, 414
Solenoid
  H field of, 238
  inductance of, 314
Spherical capacitor, 154–155, 205
Spherical coordinate system, 20–23, 72, 88, 104, 197, 243, 530
  for expressing curl, 243
  for expressing divergence, 72
  for expressing gradient, 104
  Laplacian, 197
  transformation to cartesian, 22–23
Spin
  electron, 289–292
  nuclear, 289, 292
Standing wave, 391
Standing wave ratio, 395–399, 449, 459
Stokes' theorem, 246–249
Streamlines, 46–49, 53, 54, 170–175, 229, 257
Stub matching, 461–463
Submarine communication, 372–373
Substitution of circuits, 328
Subtraction of vectors, 4
Superconductivity, 126
Superparamagnetic material, 292
Surface
  closed, 57–59, 249–250
  differential, 6, 17, 21, 59
  equipotential, 95, 101, 134, 150, 170
  gaussian, 59
  normal to, 10, 12, 45, 58, 102
  vector, 10, 58
Surface current density, 227
Surface integral, 59, 76, 250
Surface wave, 510
Surfaces
  cartesian coordinate, 6
  cylindrical coordinate, 17
  spherical coordinate, 20
Susceptibility
  electric, 142, 547–551
  magnetic, 295
Symmetry, 38, 53

Taylor's series, 68
TEM (transverse electromagnetic) wave, 355
Tesla, defined, 251
Thermal energy, 126, 136
Total reflection, 417–418, 421
Total transmission, 419–421
Toroid
  H field of, 239
  inductance of, 309
Torque, 282–287
  on differential current loop, 285
Trajectory, particle, 187
Transform-limited pulse, 430
Transformation between coordinate systems, 18–19, 22–23
Transmission coefficient, 390, 394, 415
Transmission line
  characteristic impedance, 440
  coaxial (*see* Coaxial transmission line)
  input impedance, 441
  normalized admittance, 462
  normalized impedance, 453
  planar or parallel-plate, 446–448
  two-wire, 445–446
Transmission line chart (*see* Smith chart)
Transmission line parameters, 442–448
Transmission line transients, 463–476
  current reflection diagram, 468
  initially-charged line, 472–476
  voltage reflection diagram, 466
  pulse-forming line, 474
  sign conventions on voltage and current, 468
Transmission line waves
  analogies to uniform plane waves, 438–439
  phase velocity of, 440
  phasor form of, 439
  propagation constant for, 439
  reflection at load, 440–441
  wavelength, 440
Transverse electric (TE) waves, 413
Transverse magnetic (TM) waves, 412
Transverse electromagnetic (TEM) waves, 355
Traveling wave, 354

Uncurling, 267
Uniform plane wave (*see* Plane waves, uniform)
Unit vector, 6–8, 13–14, 17, 20, 42, 45
Units, 534–539
  abbreviations for, 537–539
  conversion table for, 538

V (*see* Potential, electrostatic; Potential, scalar magnetic)
Vector(s), 2
  angle between, 10–14, 19
  component, 6, 11
  coplanar, 3
  direction of, 6, 42, 46
  equality of, 4
  Laplacian of a, 263–266
  Poynting, 367, 369
  projection of, 22
  unit (*see* Unit vector)
Vector field, 2–4, 9
Vector function, 2, 9
Vector identities, 532–533
Vector magnetic potential, 257–261, 337–341
Vector magnitude, 2, 4, 6, 8, 10–12, 42
Vector operator, 74
Vector product, 13–15
Vector projection, 12
Vector subtraction, 3
Vector surface, 11, 58
Velocity
  of charge, 121, 125
  drift, 125
  group, 425
  of light, 341, 535, 543
  phase, 353–354, 359, 410
  of propagation, 341
Volt, defined, 31, 91
Volume, differential, 6, 17, 21
Volume charge density (*see* Charge density, volume)
Volume element, 6, 17, 21
Volume integral, 36, 76

Waves
  transmission line (*see* Transmission line waves)
  uniform plane, (*see* Plane waves, uniform)
Wave equation, 340, 352

Wave impedance, 402
Waveguide (*see also specific type*)
basic operation of, 485–488
concepts of cutoff in, 493
guided modes, introduced, 487–488
modal dispersion in, 505
TE and TM modes in, 488
Wavelength, 353, 359
Wavenumber (k), 352
Wavevector (**k**), 408
Weber, defined, 251
Work (*see also* Energy), 10, 84, 87, 110

Zero reference for potential, 93

DIVERGENCE

CARTESIAN $$\nabla \cdot \mathbf{D} = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}$$

CYLINDRICAL $$\nabla \cdot \mathbf{D} = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho D_\rho) + \frac{1}{\rho}\frac{\partial D_\phi}{\partial \phi} + \frac{\partial D_z}{\partial z}$$

SPHERICAL $$\nabla \cdot \mathbf{D} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 D_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}(D_\theta \sin\theta) + \frac{1}{r\sin\theta}\frac{\partial D_\phi}{\partial \phi}$$

GRADIENT

CARTESIAN $$\nabla V = \frac{\partial V}{\partial x}\mathbf{a}_x + \frac{\partial V}{\partial y}\mathbf{a}_y + \frac{\partial V}{\partial z}\mathbf{a}_z$$

CYLINDRICAL $$\nabla V = \frac{\partial V}{\partial \rho}\mathbf{a}_\rho + \frac{1}{\rho}\frac{\partial V}{\partial \phi}\mathbf{a}_\phi + \frac{\partial V}{\partial z}\mathbf{a}_z$$

SPHERICAL $$\nabla V = \frac{\partial V}{\partial r}\mathbf{a}_r + \frac{1}{r}\frac{\partial V}{\partial \theta}\mathbf{a}_\theta + \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi}\mathbf{a}_\phi$$

CURL

CARTESIAN $$\nabla \times \mathbf{H} = \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}\right)\mathbf{a}_x + \left(\frac{\partial H_z}{\partial z} - \frac{\partial H_z}{\partial x}\right)\mathbf{a}_y + \left(\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}\right)\mathbf{a}_z$$

CYLINDRICAL $$\nabla \times \mathbf{H} = \left(\frac{1}{\rho}\frac{\partial H_z}{\partial \phi} - \frac{\partial H_\phi}{\partial z}\right)\mathbf{a}_\rho + \left(\frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho}\right)\mathbf{a}_\phi + \frac{1}{\rho}\left[\frac{\partial(\rho H_\phi)}{\partial \rho} - \frac{\partial H_\rho}{\partial \phi}\right]\mathbf{a}_z$$

SPHERICAL $$\nabla \times \mathbf{H} = \frac{1}{r\sin\theta}\left[\frac{\partial(H_\phi \sin\theta)}{\partial \theta} - \frac{\partial H_\theta}{\partial \phi}\right]\mathbf{a}_r + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial H_r}{\partial \phi} - \frac{\partial(rH_\phi)}{\partial r}\right]\mathbf{a}_\theta + \frac{1}{r}\left[\frac{\partial(rH_\theta)}{\partial r} - \frac{\partial H_r}{\partial \theta}\right]\mathbf{a}_\phi$$

LAPLACIAN

CARTESIAN $$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}$$

CYLINDRICAL $$\nabla^2 V = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho\frac{\partial V}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 V}{\partial \phi^2} + \frac{\partial^2 V}{\partial z^2}$$

SPHERICAL $$\nabla^2 V = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial V}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 V}{\partial \phi^2}$$